



# مرور ساختار یافته و متا- آنالیز، مفاهیم، کاربردها و محاسبات

تالیف: دکتر علی اکبر حقدوست  
استادیار آمار و اپیدمیولوژی دانشگاه علوم  
پزشکی کرمان

تحت حمایت موسسه پژوهشگران بدون مرز

فهرست مطالب

۶	مقدمه
۱۰	فصل اول
۱۰	۱. تعریف دقیق موضوع تحقیق و سوال پژوهش
۱۱	۱-۱. جامعه مورد مطالعه
۱۲	۲-۱. متغیر مستقل مورد مطالعه
۱۳	۳-۱. متغیر وابسته مورد مطالعه
۱۳	۴-۱. انواع مطالعات
۱۴	۵-۱. اهمیت سوال پژوهش در مراحل بعدی
۱۸	فصل دوم
۱۸	۲. جستجوی مطالعات مرتبط به سوال پژوهش
۱۹	۱-۲. حساسیت و اختصاصی بودن جستجو
۲۰	۲-۲. بانک های اطلاعاتی الکترونیک
۲۱	مدلاین و Embase
۲۳	بانک اطلاعات کارآزمایی های بالینی ککراین (CENTRAL)
۲۴	۳-۲. جستجوی دستی منابع غیر الکترونیکی
۲۵	۴-۲. جستجوی سایر مقالات مروری مرتبط
۲۵	۵-۲. بانکهای اطلاعات کاغذی
۲۵	۶-۲. شناسایی منابع چاپ نشده
۲۶	۷-۲. تدوین استراتژی جستجو
۳۰	۸-۲. سوگیری در انتشار (publication bias)
۳۴	فصل سوم
۳۴	۳. سنجش کیفیت مطالعات
۳۵	۱-۳. پایایی
۳۶	۲-۳. اعتبار
۳۹	۳-۳. سنجش کیفیت مطالعات

۴۲	۴-۳. شیوه سنجش کیفیت مطالعات در عمل
۴۴	۵-۳. رابطه بین احتمال انتشار مقاله و کیفیت مطالعه
۴۷	<b>فصل چهارم</b>
۴۷	۴. استخراج نتایج
۴۸	۱-۴. قسمت‌های اصلی فرم جمع‌آوری اطلاعات
۴۹	اطلاعات مربوط به مطالعه و فرد جمع‌آوری‌کننده اطلاعات
۴۹	اطلاعات مربوط به شرایط‌گزینش مطالعه
۵۰	مشخصات مطالعه
۵۱	۲-۴. فرم‌های الکترونیک و کاغذی
۵۳	۳-۴. انجام پیش‌مطالعه برای شناسایی مشکلات فرمها
۵۴	<b>فصل پنجم</b>
۵۴	۵. آماده‌سازی اطلاعات برای متاآنالیز
۵۵	۱-۵. شاخص و پارامتر
۵۶	۲-۵. شاخص‌های قابل اندازه‌گیری در مطالعات توصیفی
۵۷	۳-۵. شاخص‌های قابل محاسبه در مطالعات تحلیلی
۵۷	معنی‌داری آماری
۵۸	شدت اثر
۶۰	شاخص‌های نشان‌دهنده شدت اثر
۷۲	۴-۵. مقیاس شاخصها در متاآنالیز
۷۲	۵-۵. شاخص‌های خام و تصحیح شده
۷۴	۶-۵. آخرین حربه برای متاآنالیز مطالعات غیرمتجانس
۷۵	<b>فصل ششم</b>
۷۵	۶. مفاهیم اولیه متاآنالیز
۷۵	۱-۶. متاآنالیز در مطالعات توصیفی
۷۷	۲-۶. متاآنالیز در مطالعات تحلیلی
۷۷	اهداف متاآنالیز مطالعات تحلیلی
۷۸	گسترده‌گی رابطه مورد بررسی
۸۰	۳-۶. نقش متاآنالیز

۸۲	۴-۶. موارد نابجا استفاده از متاآنالیز
۸۳	۵-۶. مقدمات آماری متاآنالیز
۸۵	۶-۶. مدل فیکس در مقایسه با مدل تصادفی
۸۹	۷-۶. روشهای آماری رایج برای ترکیب نمودن شاخصها
۸۹	روش معکوس واریانس (Inverse Variance Weighted Method)
۹۵	۸-۶. روشهای آماری رایج برای ترکیب نمودن شاخصها در مدل‌های تصادفی
۹۷	روش Der Simonian and Laird
۹۷	روش متا-رگرسیون
۱۰۱	۹-۶. مفهوم عدم تجانس (Heterogeneity)
۱۰۲	نحوه اندازه گیری میزان عدم تجانس
۱۰۳	شیوه برخورد با عدم تجانس آماری مطالعات
۱۰۷	۷. نمودار های اصلی در متاآنالیز و آزمونهای آماری مرتبط
۱۰۷	7-1. نمودار انباشت (Forest Plot)
۱۱۲	۲-۷. نمودار قیفی (Funnel Plot) برای بررسی شانس بروز سوگیری در انتشار نتایج
۱۱۷	۳-۷. آزمونهای آماری مرتبط به سوگیری در انتشار نتایج
۱۱۷	آزمون آماری Begg and Manzumdar
۱۱۸	آزمون آماری Egger regression asymmetry test
۱۲۰	۴-۷. نمودار Galbraith
۱۲۲	۵-۷. نمودار L'Abbe
۱۲۵	۸. موضوعات ویژه در متاآنالیز
۱۲۵	۱-۸. متاآنالیز تجمعی Cumulative Meta-Analysis
۱۲۷	۲-۸. متا-آنالیز P-values
۱۲۸	فرمول فیشر
۱۲۹	روش Edgington
۱۳۰	۳-۸. ترکیب اطلاعات تک تک نمونهها مطالعات (IPD)
۱۳۲	۴-۸. تحلیل حساسیت (sensitivity analysis)
۱۳۵	۵-۸. تاثیر نمره کیفیت مطالعات در وزن داده شده در متا-آنالیز
۱۳۸	۹. دستورات برنامه STATA برای انجام متا-آنالیز
۱۳۹	9-1. نحوه اجرای دستورات متا-آنالیز

۱۴۱	۲-۹. دستورات اصلی متا-آنالیز
۱۴۲	دستور meta
۱۴۶	دستور metan
۱۴۷	دستور metacum
۱۴۹	دستور metabias
۱۵۲	دستور metatrim
۱۵۲	دستور labbe
۱۵۳	دستور funnel
۱۵۳	دستور metap
۱۵۴	دستور metainf
۱۵۶	دستور metatrim
۱۵۸	۱۰. منابع

## مقدمه

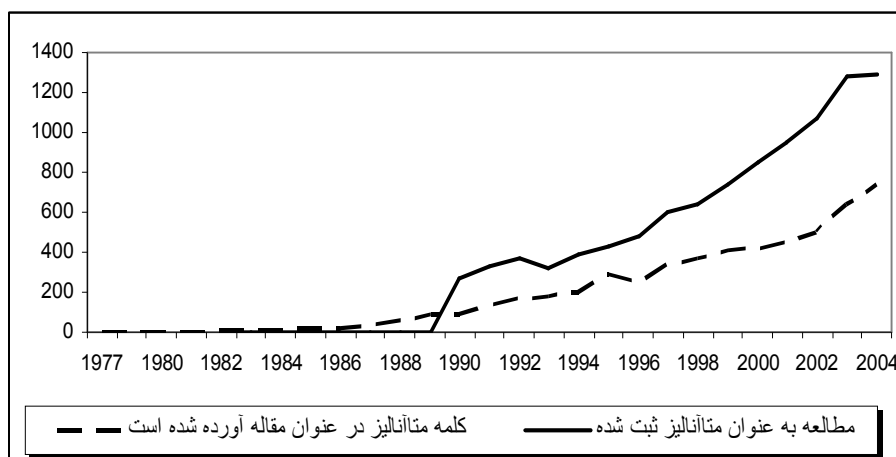
بدلیل رشد بسیار سریع علمی، ارائه دهندگان خدمات بهداشتی درمانی، محققین و سیاستگذاران سلامت با حجم عظیمی از اطلاعات روبرو هستند. در مواجهه با این انفجار علمی و برای استخراج سریع و دقیق اطلاعات، لازم است افراد بصورت ساختار یافته (systematic review) منابع موجود جستجو نمایند. جستجوی ساختار یافته کمک میکند تا سوگیریها به حداقل ممکن برسد و خطاهای تصادفی نیز کاهش یابد. بدین دلیل اطلاعات استخراج شده، بیشترین پایایی (reliability) و اعتبار (validity) را خواهد داشت.

از طرف دیگر، متاآنالیز (meta-analysis) عبارت است از بکارگیری روش های آماری خاص برای خلاصه نمودن نتایج مطالعات مستقل برای یافتن دقیقترین شکل ارتباط بین متغیرهای مورد بررسی. این روشهای آماری کمک میکنند تا جمع بندی اطلاعات مقالات مختلف و خلاصه نمودن آنها بصورت عینی (objective) صورت پذیرند و نظرات شخصی تاثیر خاص در این فرآیند نداشته باشد.

تا قبل از دهه ۱۹۷۰ عمده مقالات مروری توسط افراد صاحب نظر نگارش میشد و در بهترین شکل توسط یک گروه مشخص ویرایش نهایی انجام میشد. در دهه ۱۹۷۰ به تدریج محققین علوم اجتماعی و روانشناسی خصوصیات یک جستجوی ساختار یافته را شرح دادند و دلایل بهتر بودن این نوع جستجو را در کاهش خطاهای تصادفی بیان

## جستجوی ساختاریافته و متاآنالیز

نموند. در اواخر دهه ۱۹۸۰ بصورت علمی ایرادات مقالات مروری رایج توسط چندین محقق شرح داده شد و بدین شکل گرایش جامعه عملی به سمت جستجو و مرور ساختار یافته افزایش یافت. در دهه ۱۹۹۰ با ایجاد گروههای تخصصی، جستجوی ساختار یافته در دنیا رایج شد و با ایجاد بانک های اطلاعاتی خاص مانند ککراین (Cochrane) شتاب آن فزونی یافت.



نمودار ۱: تعدیل کفالاتم دلاین به ۴ عذ وان مطالعه ات متاآذ الیز واراندر عذ وان ذ ود کلم ۴ متاآذ الیز را دارناولذ یذ مقاله ۴ درس الوالااااا ده ک ۴ اص ول انج ام متاآذ الیز را ش رح دانامذ ت ( Smith ML et. ) (al.)

اگرچه تا کنون کتاب و مقالات بسیار زیادی در خصوص متدولوژی جستجو ساختار یافته و متاآنالیز به چاپ رسیده و سرعت کشف و ارائه روشهای آماری جدید در این زمینه قابل توجه است، با این وجود تاکنون در منابع فارسی به این مطلب چندان پرداخته نشده است. به همین دلیل کتاب حاضر تدوین شده است تا به علاقمندان کمک نماید با متدولوژی بررسی ساختار یافته متون بهتر آشنا شده و روشهای متاآنالیز را فرا گیرند. در این کتاب کمتر به مباحث آماری پرداخته شده است و نحوه

تنظیم آن بگونه ای است که افراد با کمترین آگاهی آماری نیز بتوانند از مطالب این کتاب بهره ببرند.

از آنجایی که احتمالاً خوانندگان این کتاب طیف وسیعی را تشکیل می‌دهند (پزشکان و متخصصین بالینی تا کارشناسان و متخصصین آمار و اپیدمیولوژی)، لذا سعی شده است که محتوای کتاب بگونه‌ای تنظیم شود که هر گروه به راحتی و بدون اتلاف وقت بتوانند به مطالب مورد نیاز خود دست یابند. بدین جهت قسمتهای پایه و مفاهیم اولیه در فصل‌های خاصی متمرکز شده اند که افراد آشنا با مبانی بتوانند با مرور سریع این فصل-ها، وقت کمتری را صرف آن بنمایند. همچنین مطالب تخصصی‌تر که جنبه آماری آن بیشتر است در کادرهای خاکستری گذاشته شده است و افراد غیر متخصص در آمار می‌توانند این قسمت‌ها را نادیده گرفته و بدون گم نمودن سر نخ، مطالب کتاب را دنبال نمایند.

این کتاب دو هدف اصلی را دنبال می‌کند:

۱. آشنا نمودن خوانندگان با مفاهیم اصلی جستجو ساختاریافته و نقد مقالات متاآنالیز؛ برای درک بهتر و عمیق‌تر مقالات چاپ شده در این زمینه

۲. توانمند نمودن خوانندگان برای انجام جستجوی ساختار یافته و نشر مروری و متاآنالیزی

مسلماً سطح دوم بسیار عمیقتر از سطح اول است. لذا توصیه میشود علاقمندان برای دستیابی به این سطح از فراگیری، ضمن مطالعه این کتاب درکارگاههایی که در این راستا برگزار میشود نیز شرکت نمایند.

برای محاسبات در این کتاب از نرم افزار Stata استفاده شده است که توانمندی بسیار زیادی در انجام متاآنالیز دارد. همچنین بصورت گذرا سایر نرم افزارهای قابل استفاده، در مراحل مختلف معرفی می‌شوند.



بر اساس تقسیم بندی سایت ککراین (Cochrane) برای نگارش یک مقاله متاآنالیز حداقل شش مرحله به شرح زیر نیاز است:

۱. تعریف دقیق موضوع و سوال پژوهش
  ۲. انتخاب اولیه مقالات در قالب یک بررسی ساختاریافته اولیه
  ۳. بررسی اعتبار علمی مقالات و منابع
  ۴. استخراج اطلاعات
  ۵. تحلیل آماری و ارائه آن
  ۶. تفسیر و بحث و نتایج
- بر این اساس، عناوین فصول مختلف این کتاب نیز به همین شکل نظم یافته است.

## فصل اول

### ۱. تعریف دقیق موضوع تحقیق و سوال پژوهش

در این فصل شما با جایگاه سوال پژوهش در مطالعات مروری و متاآنالیز آشنا خواهید شد و در می‌یابید که چگونه سوال پژوهش می‌توانند در مراحل بعدی کار تاثیرگذار باشد. همچنین با ویژگیهای سوال صحیح آشنا خواهید شد.

هدف از مرور ساختاریافته و انجام متاآنالیز پاسخ به یک سوال خاص مبتنی بر نتایج تحقیقات مختلف است. بر این اساس باید نتایج تحقیقات مختلف قابل ترکیب نمودن باشند؛ چراکه در غیر این صورت، بدلیل تفاوت‌های اساسی بین تحقیقات، نتیجه‌گیری دقیقی صورت نخواهد گرفت.

پرواضح است که اگر سوال مورد پژوهش بدقت طراحی نشود در گزینش مقالات، مشکلات عدیده‌ای ایجاد خواهد شد. بنابراین قبل از شروع هر کاری، محقق بایست به دقت دامنه تحقیق را مشخص و سوال تحقیق خود را بیان نماید بعنوان مثال فرض نماید که می‌خواهد در خصوص رابطه بین بیماری عروق کرونر و ورزش یک بررسی ساختاریافته انجام داده و نتایج را متاآنالیز نماید. از آنجایی که هدف مطالعه خلاصه نمودن نتایج تحقیقات بصورت کمی (Quantitative) است، لذا بایست متغیرهای اندازه‌گیری شده و جامعه‌های مورد بررسی قابل ترکیب شدن باشند. بیماری عروق کرونر از تنگی‌های بی‌علامت شروع شده و بعد از علامت‌دار شدن می‌تواند

نماهای بالینی متفاوتی داشته باشد که مهمترین آنها آنژین صدری پایدار، آنژین صدری ناپایدار، حمله حاد قلبی، سکته و حتی مرگ ناگهانی می‌باشند. همچنین ورزش و فعالیت‌های بدنی نیز طیف وسیعی را تشکیل می‌دهند. آنچه واضح است این است که تعریف بیماری عروق کرونر و ورزش در تحقیقات مختلف یکسان نیست. لذا جمع‌بندی این نتایج ممکن است خالی از اشکال نباشد. این شما هستید که با توجه به موضوع تحقیق بایست تصمیم بگیرید تا چه حد تعاریف این متغیرها را باز نموده و امکان ورود مطالعات بیشتر را فراهم آورید.

بعد از این مثال ساده سعی می‌شود در ادامه، حیطه‌های مختلفی را که سوال پژوهش بایست به آنها بپردازد بیان گردد.

### ۱-۱. جامعه مورد مطالعه

محقق بایست به دقت خصوصیات با اهمیت جامعه مورد مطالعه را تعیین و شرح دهد. شدت بیماری و معیارهای تشخیصی آن در بسیاری موارد با اهمیت هستند. مسلماً تاثیر داروها در مراحل مختلف بیماری یکسان نیست و معیارهای تشخیصی معمولاً دقت یکسانی در تشخیص ندارند. همچنین یک بیماری ممکن است تظاهرات بالینی متفاوت داشته‌باشد و درمان‌های شناخته شده تاثیرات متفاوت بر این تظاهرات داشته‌باشند. به همین دلیل جمع‌زدن نتایج تحقیقات مختلف در بعضی موارد ممکن است دارای ایراد باشد.

از طرف دیگر متغیرهای زمینه‌ای مانند سن، جنس، نژاد، وضعیت تحصیل و اجتماعی نمونه‌ها نیز ممکن است بر پاسخ افراد تاثیرات قابل ملاحظه‌ای بگذارد. بعنوان مثال تاثیر ماموگرافی در تشخیص مراحل اولیه بدخیمی‌های پستان بستگی به سن دارد. لذا ترکیب نتایج مختلف

تحقیقات مختلف با جمعیت مورد مطالعه متفاوت از نظر سنی خالی از اشکال نیست.

از طرف دیگر، محدود نمودن معیارهای گزینش مطالعات اولاً باعث کم شدن تعداد مطالعات مورد پذیرش می‌شود و از طرف دیگر تعمیم پذیری نتایج به کل جامعه را کاهش می‌دهد. به همین دلیل محقق بایست در تعریف جامعه هدف خود را با دقت زیاد و مبتنی بر ادله قابل دفاع بنا نماید.

### ۱-۲. متغیر مستقل مورد مطالعه

این متغیر عمدتاً نوع تداخلی (Intervention) است که بررسی می‌شود. بعنوان مثال اگر هدف بررسی اثر مصرف سیگار بر بدخیمی مثانه است تعریف سیگاری بودن و اندازه گیری شدت آن بسیار با اهمیت است.

در کارآزمایی‌های بالینی بر روی داروها، مهمترین فاکتورهای قابل توجه دز دارو، شیوه مصرف، مدت مصرف، داروی مورد مقایسه (درمان استاندارد و یا پلاسیبو) و حتی کیفیت ساخت دارو می‌باشد.

البته در مقایسه با سایر مداخلات، تعریف متغیر مستقل در کارآزمایی بالینی بر روی تاثیرات داروها آسان‌تر می‌باشد. بعنوان مثال ارائه تعریف مناسب برای مداخلاتی که به شکل آموزش به جامعه داده می‌شود گاه بسیار دشوار می‌باشد، چرا که محتوی شیوه ارائه آموزش از یک تحقیق به تحقیق دیگر می‌تواند بسیار متفاوت باشد.

### ۳-۱. متغیر وابسته مورد مطالعه

در ابتدای بحث پیچیدگی‌های تعریف بیماری‌های عروق کرونر بیان شد. در عمل ارائه تعریف دقیق و در عین حال گسترده از متغیر وابسته که عمدتاً نتیجه مداخله (outcome) می‌باشد بسیار دشوار است. بعنوان یک روال کلی در مطالعات مروری ککراین، متغیر وابسته شامل تمامی حالت‌های ممکن از یک پدیده می‌باشد که افراد صاحب نظر در تصمیمات خود برای گزینش بهترین روش‌های مقابله با آن پدیده، آنها را مد نظر قرار می‌دهند. بعنوان مثال از آنجایی که کلیه تظاهرات بالینی نارسایی عروق کرونر بصورت بالقوه می‌توانند خطرناک باشند در این تعریف بایست وارد شوند.

اگرچه این تعریف به ظاهر بسیاری از مشکلات را حل می‌کند ولی در عمل می‌تواند مشکلات جدیدی نیز بوجود آورد و باعث ناهمگونی بسیار زیاد مطالعات گزینش شده شود. بعنوان مثال برای بررسی اثرات یک داروی ضد سرطان، بعضی مطالعات ممکن است میزان بقا را بعنوان متغیر وابسته خود در نظر گرفته‌باشند و بعضی دیگر عوارض بیماری و درمان را و عده ای کیفیت زندگی بیماران را. پرواضح است که وارد نمودن کل این مطالعات و جمع بندی کمی نتایج مشکل‌آفرین است.

### ۴-۱. انواع مطالعات

برای بررسی رابطه متغیرها می‌توان مطالعات مورد-شاهد، کوهرت و یا مداخله‌ای انجام داد. حال این سوال مطرح است که آیا می‌توان در متاآنالیز نتایج این مطالعات را روی هم ریخت و جمع بندی نمود. مسلماً دقت این مطالعات یکسان نیست و خطاهای تصادفی و بخصوص سرگیری‌ها (Bias) در مطالعات مداخله‌ای اکثراً کمتر است.

اما در عوض در بعضی موارد، مطالعات کوهرت به جهت مدت زمان بیشتر پیگیری ارزشمند می‌باشد. همچنین شاید برای شناسایی عوامل خطر بیماری‌های نادر جز مطالعات مورد-شاهدی جایگزین دیگری وجود نداشته‌باشد.

حتی مطالعات مداخله‌ای نیز خود انواع مختلفی دارند و کارآزمایی بالینی (clinical trial) یک نوع آنها می‌باشد. بعنوان مثال ثابت شده است که مطالعات مداخله‌ای بدون رعایت تصادفی وارد نمودن نمونه‌ها در گروه‌ها (Randomization) معمولاً اثرات مداخله‌ها را بیشتر از حد واقعی گزارش می‌کنند. همچنین خطاهای کارآزمایی‌های بالینی به روش کلاسیک (داری گروه‌های مداخله و پلاسیبو) با کارآزمایی‌هایی که به روش Cross-over انجام می‌شود یکسان نیست.

با توجه به این نکات، بسیار اهمیت دارد که محقق در ابتدا مشخص نماید که چه نوع مطالعاتی را گزینش و نتایج آنها را وارد متاآنالیز خواهد نمود. اگر چه تمایل عمومی، محدود نمودن مطالعات به کارآزمایی‌های بالینی است ولی در بعضی موارد، همانند وجود مشکلات اخلاقی در انجام مطالعات مداخله‌ای، حجم نمونه کم و مشکلات اجرایی در آنها، وارد نمودن نتایج سایر انواع مطالعات نیز پذیرفته است.

### ۱-۵. اهمیت سوال پژوهش در مراحل بعدی

مسلماً استراتژی جستجوی منابع اطلاعاتی وابسته به سوال پژوهش است. کلید واژه‌های انتخابی باید به شکلی باشند که موضوع مورد پژوهش را به خوبی پوشش دهند. اگر قصد وارد نمودن کارآزمایی‌های بالینی است بایست

بیشترین توجه به بانک های اطلاعاتی اختصاصی برای این نوع مطالعات باشد. همچنین گذاشتن کلید واژه های زیاد در کنار یکدیگر با ترکیب "and" باعث محدود شدن دامنه جستجو و حذف تعدادی از مقالات مرتبط می گردد. در مقابل، باز گذاشتن دامنه جستجو باعث گزینش تعداد زیاد مقالات غیر مرتبط می شود. بنابر این بهترین ترکیب کلید واژه ها بر اساس سوال پژوهش شکل می گیرد.

بعد از جستجوی بانک های اطلاعاتی و در مرحله استخراج نتایج نیز، سوال پژوهش بسیار با اهمیت است. بعنوان مثال اگر در مطالعه ای سن نمونه ها با اهمیت تشخیص داده شده باشد بایست در ثبت اطلاعات، این متغیر نیز بدقت استخراج شود.

مرحله تحلیل نتایج نیز وابسته به سوال مورد پژوهش است، بعنوان مثال در ارزیابی عوامل موثر در ناهمگون بودن نتایج مطالعات (Heterogeneity)<sup>۱</sup> فاکتورهایی همچون تعریف متغیرهای مستقل و وابسته، جمعیت مورد مطالعه و نوع مطالعات می تواند مهم بوده و در تحلیلها وارد شوند.

سوالات گسترده معمولاً حجم کار را زیاد نموده و باعث ناهمگونی بیشتر نتایج مطالعات می شود در عوض سوالات محدود باعث کم شدن تعداد مطالعات گزینش شده و کاهش تعمیم پذیری نتایج به کل جامعه می شوند.

استراتژی جستجوی منابع برای بررسی اثرات فاکتورهای ایجاد کننده ناهمگونی نیز می تواند وابسته به شکل سوال پژوهش باشد. در بعضی موارد ممکن است محقق از ابتدا سوال را گسترده طراحی نموده و سپس جستجوی خود را بصورت باز انجام دهد؛ ولی در بعضی موارد ممکن

---

<sup>۱</sup> منظور عواملی است که باعث به وجود آمدن تفاوت در نتایج تحقیقات مختلف می شود. به واقع متاآنالیز دو عملکرد متفاوت دارد، عملکرد اول آن ترکیب نمودن نتایج تحقیقات مختلف است و عملکرد دوم آن شناسایی عوامل ایجاد کننده ناهمگونی نتایج مطالعات مختلف

است ترجیح برآن باشد که چند سوال کاملاً محدود طراحی شده و برای هر سوال بصورت جداگانه جستجوی صورت پذیرد. مثلاً در بررسی بیماری عروق کرونر ترکیب نتایج چند جستجوی جداگانه بر روی تظاهرات بالینی متفاوت این بیماری، احتمالاً روش بهتری است؛ در عوض در بررسی تأثیر مصرف میوه بر یبوست، جستجوی جداگانه میوه‌های مختلف کار سنگین است و بکار بردن کلید واژه‌های عمومی همچون "میوه" مناسبتر می‌باشد. البته شرط استفاده از این روش آن است که هدف اصلی پژوهش پاسخ به یک سوال کلی باشد و مثلاً ماده غذایی و یا ویتامین خاصی مدنظر نباشد. با این نگاه در بعضی مطالعات ممکن است بتوان نتایج تحقیقات بر روی خواص میوه‌هایی مثلاً سیب و پرتقال را ترکیب و در بعضی مطالعات، ترکیب نمودن نتایج آنها کاملاً غیر قابل توجیه باشند.

آخرین نکته، امکان تغییر سوال پژوهش در حین مطالعه می‌باشد. در بین کار ممکن است محقق به این جمع‌بندی برسد که سوال مورد تحقیق کامل و یا دقیق نبوده و نیاز به تصحیح دارد. اگر چه با یک توجیه قوی این امکان وجود دارد ولی بایست توجه داشت که سوالات تغییر یافته در حین کار عموماً دارای سوگیری بیشتری در مقایسه با سوالات از پیش تعریف شده هستند و به همین دلیل قبل از هر تغییر بایست محقق به سوالات زیر به دقت پاسخ دهد.

۱. انگیزه اصلی تصحیح چیست؟

۲. آیا تحت تأثیر نتایج بعضی از مطالعات قرار گرفته-اید یا در حین کار متوجه شده‌اید که در ابتدا به کلیه جوانب و عوامل موثر اشراف نداشته‌اید؟

۳. آیا بعد از تغییر سوال، استراتژی اخذ شده برای جستجوی بانک‌های اطلاعاتی کماکان قابل قبول هستند؟



۴. آیا بعد از تغییر سوال، شیوه استخراج کماکان قابل قبول میباشد؟

## فصل دوم

### ۲. جستجوی مطالعات مرتبط به سوال پژوهش

---

این فصل به شما کمک می‌کند تا به شکل صحیحی استراتژی جستجوی خود را تنظیم نموده و از منابع موجود استفاده نمایید. این فصل شما را تنها با اصول جستجوی بانکهای اطلاعاتی الکترونیک و همچنین غیر الکترونیک آشنا می‌کند و در صورتی که شما با نحوه جستجو آشنا نیستید باید از منابع آموزشی مربوطه استفاده نمایید.

جستجو و یافتن مطالعات مرتبط به موضوع مورد مطالعه یکی از حساسترین و وقتگیرترین مراحل است. در اکثر جستجوهای ساختاریافته برای بررسی اثر بخشی مداخلات کارآزمایی بالینی بعنوان تنها شکل مطالعات پذیرفته شده گزینش می‌شوند. ولی باید توجه داشت که در مواردی خاص بایست سایر انواع مطالعات را نیز وارد نمود. حتی گاهی اوقات مطالعات توصیفی مقطعی نیز در جستجوهای ساختاریافته وارد می‌شوند. به عنوان مثال اگر هدف یک مطالعه تعیین میزان پرفشاری خون در کشور ایران باشد، محقق باید تنها به دنبال مطالعاتی باشد که هدف آنها توصیف فشار خون افراد جامعه بوده است.

مطالعات نشان داده‌اند که یک جستجو سریع مدلاین تنها ۳۰ الی ۸۰ درصد کارآزمایی بالینی مرتبط را شناسایی می‌کند. بنابراین لازم است از همان ابتدا

استراتژی دقیقی برای جستجوی منابع و بانک های اطلاعاتی تعیین شود. بعنوان مثال جستجوی بانک اطلاعاتی مدلاین بعنوان تنها منبع اطلاعات می تواند باعث سوگیری شود (Publication bias and language bias). چرا که معمولاً مقالات ایندکس شده در مدلاین به زبان انگلیسی هستند و در این مقالات احتمال گزارش شدن یک اختلاف معنی دار آماری بیشتر از مقالات غیر انگلیسی است.

در این فصل سعی می شود به مفاهیم اصلی شیوه جستجوی ساختار یافته اشاره و نکات ارزشمندی برای بالا بردن دقت جستجو بیان می شود.

### ۱-۲. حساسیت و اختصاصی بودن جستجو

حساسیت (sensitivity) یک جستجو بیان کننده آن است که چند درصد مقالات مرتبط، کشف و گزینش شده اند. مسلماً محققین تمایل دارند تا حساسیت جستجوی خود را تا حد امکان بالا ببرند. به همین دلیل توصیه می شود چندین بانک اطلاعات همزمان با کلیه واژه های نسبتاً عام جستجو شود و مقالات به زبانهای مختلف بررسی شوند. همچنین با افزایش یافتن تعداد سالهای مورد بررسی، احتمالاً حساسیت جستجو بهتر میشود.

در مقابل اختصاصی بودن (specificity) مبین آن است که تا چه میزان مقالات غیر مرتبط گزینش شده اند. مسلماً وارد شدن تعداد بسیار زیاد مقالات غیر مرتبط باعث افزایش حجم کار و حتی کاهش دقت در مطالعه آنها میشود.

به سادگی قابل تصور است که دو شاخص فوق نسبت معکوس دارند یعنی با استفاده از کلید واژه های

عامتر حساسیت جستجو افزایش می‌یابد ولی در مقابل بدلیل بیشتر شدن مقالات غیر مرتبط، اختصاصی بودن جستجو کاهش می‌یابد.

به هر شکل محقق بایست در ابتدا با دقت و صرف زمان کافی سعی نماید بهترین استراتژی جستجو را تعیین نماید تا میزان اختصاصی بودن و حساسیت جستجوی خود را به حداکثر برساند.

### ۲-۲. بانک های اطلاعاتی الکترونیک

معمولاً اکثر جستجوها از منابع الکترونیک شروع می‌شود. خوشبختانه بانکهای اطلاعاتی الکترونیک در حیطه علوم پزشکی به خوبی گسترش پیدا کرده‌اند و به همین دلیل جستجو برای محققین آسانتر شده است. مهمترین بانکهای اطلاعاتی مدلاین (Medline) و Embase می‌باشند. البته بجز این دو بانک دهها بانک اطلاعاتی دیگر وجود دارد که بسته به موضوع محقق باید مهمترین آنها را انتخاب و جستجو نماید. بعضی از مهمترین این بانک ها عبارتند از:

- Australasian Medical Index
- Chinese Biomedical Literature Database
- Japan Information of Science and Technology File on Science
- Latin American Caribbean Health Science Literature (LILACS)
- Cumulative Index of Nursing and Allied Health (CINAHL)
- AIDSLINE
- CAB Health
- SciSearch

همچنین ککراین به تازگی بانک اختصاصی برای کارآزمایی‌های بالینی تشکیل داده است (CENTRAL)، که توضیحات بیشتر آن در ادامه خواهد آمد. همچنین علاقمندان برای کسب اطلاعات بیشتر در خصوص انواع بانک‌های اطلاعاتی و محتویات آنها می‌توانند به آدرس اینترنتی زیر مراجعه نمایند:

<http://www.dialog.com>

با توجه به اهمیت بانک‌های مدلاین، CENTRAL، Embase در ادامه خصوصیات آنها بصورت مبسوط‌تر ارائه خواهد شد.

### مدلاین و Embase

ایندکس مدیکوس (Index Medicus) توسط کتابخانه ملی پزشکی آمریکا تهیه می‌شود و فرم الکترونیکی آن مدلاین می‌باشد. در مقابل Excerpta Medica توسط Elsevier تهیه می‌شود و فرم الکترونیک آن Embase می‌باشد. مدلاین، ۴۶۰۰ ژورنال را از آمریکا و ۷۰ کشور دنیا پوشش می‌دهد و تا فوریه ۲۰۰۵ دارای بیش از ۱۴ میلیون رکورد بوده است. این بانک از سال ۱۹۶۶ تشکیل شده ولی اخیراً تعدادی از منابع قبل از ۱۹۶۶ نیز به آن اضافه شده است. در حال حاضر امکان جستجوی مجانی این بانک از طریق آدرس <http://www.ncbi.nlm.nih.gov> مهیا می‌باشد.

Embase مجلات اروپایی را بیشتر از مدلاین پوشش می‌دهد و تقریباً ۴۰۰۰ عنوان مجله را از ۷۰ کشور جهان شامل می‌شود و تا آخر سال ۲۰۰۴ میلادی بالغ بر ۱۱ میلیون رکورد داشته است.

هم پوششی این دو بانک اطلاعاتی تقریباً ۳۴٪ می‌باشد ولی بسته به موضوع جستجو این همپوشانی معمولاً بین ۱۰ الی ۷۴ درصد گزارش شده است.

شیوه جستجو مدلاین و Embase تفاوت‌هایی دارد به همین دلیل برای جستجو آنها بایست محقق کلید واژه‌ها و استراتژی‌های متفاوت برگزینند. بعنوان مثال از سال ۱۹۹۱ به بعد، امکان جستجوی جدیدی به مدلاین اضافه شده است و محقق می‌تواند عبارت (Randomized-Controlled-Trial) را در نوع انتشار (Publication type) وارد نماید و بدین شکل محدوده جستجو خود را مشخص نماید، در حالی که این امکان برای رکوردهای قبل از ۱۹۹۱ در مدلاین و همچنین در کل بانک Embase وجود ندارد. با توجه به مجانی بودن جستجوی مدلاین و همچنین تحت پوشش قرار گرفتن تعداد زیادی از مجلات سایر بانکها در آن، معمولاً اکثر محققین جستجوی خود را با مدلاین شروع می‌کنند. همچنین کتابخانه ملی پزشکی آمریکا اخیراً بانک جدیدی بنام Gateway راه اندازی نموده است که در آن می‌توان همزمان بانکهای زیر را جستجو نمود:

- 1- Pubmed
- 2- old medline
- 3- locator puls
- 4- Medline plus
- 5- Dirline
- 6- AIDS Meetings
- 7- Health Services Research Meetings
- 8- Space Life Sciences Meetings
- 9- HSRProj

آدرس Gateway در اینترنت  
<http://gateway.nlm.nih.gov/gw/cmd> می‌باشد.

باید مجدد بیان نمود که تکیه کردن صرف به منابع الکترونیکی برای جستجو ناکافی بوده و ممکن است باعث کاهش حساسیت جستجو شود (مراجعه شود به بخش ۱-۲).

جدول ۱: مقایسه بین تعداد کارآزمایی‌های بالینی وارده شده در دو متاآنالیز و تعدادی که در جستجوی مدلاین انتخاب شده بوده‌اند.

موضوع متاآنالیز	تعداد مقالات وارد شده در بانک اطلاعات prenatal	تعداد مقالات یافت شده در مدلاین جستجوی فرد باتجربه	تعداد مقالات یافت شده در مدلاین جستجوی فرد کم تجربه
Neonatal hyperbilirubinemia	۸۸	۲۸	۱۷
Interaventricular hemorrhage	۲۹	۱۹	۱۱

Source: Dickersin et. al. (1985) +page 52

## بانک اطلاعات کارآزمایی‌های بالینی ککراین (CENTRAL)

این بانک که اسم کامل آن The Cochrane Control Trial Central Register می‌باشد، در حال حاضر کاملترین سایت در نوع خود می‌باشد و تا پایان سال ۲۰۰۴ بالغ بر نیم میلیون رکورد بوده است. این سایت مطالعات کارآزمایی بالینی و مقالات مروری ککراین را پوشش می‌دهد و در آن مقالات به زبانهای مختلف یافت می‌شود.

مراکز ککراین در آمریکا، انگلستان، استرالیا، چین، برزیل و چند کشور دیگر مسئول به روز نمودن اطلاعات این بانک هستند. پرونیکل جستجوی آنها یکسان می‌باشد و راهنمای جستجوی این بانک اطلاعاتی در سایت <http://www.cochrane.us/manage.htm> قابل دسترس می‌باشد.

### ۲-۳. جستجوی دستی منابع غیر الکترونیکی

بعد از جستجوی منابع الکترونیک معمولاً نیاز است تا محقق به صورت دستی عناوین مقالات مجلات مرتبط را به صورت دستی مرور نماید. این نوع جستجو اولاً کمک می‌کند تا مقالاتی که در جستجوی الکترونیک انتخاب نشده‌اند وارد مطالعه شوند و ثانیاً بعضی منابع مانند supplements که در مدلاین ایندکس شده‌اند نیز بررسی شود. همچنین خلاصه مقالات کنفرانسها **conference proceeding** از منابع بسیار مهمی است که محقق باید به صورت دستی آنها را جستجو کند. به صورت متوسط ۱۰ درصد منابع در مقالات مروری ککراین از خلاصه مقالات کنفرانسها انتخاب می‌شوند و تحقیقات نشان داده‌اند که حدود ۵۰ درصد خلاصه‌های ارائه شده در کنفرانسها هرگز به مقاله تبدیل نمی‌شوند و لذا جستجو در مدلاین و یا Embase نمی‌تواند آنها را کشف نماید. به عنوان راهنما برای جستجو دستی توصیه می‌شود مطالعه مندرج در آدرس زیر مطالعه شود:

<http://www.cochrane.us/hsmain.htm>

ولی بایست توجه داشت که تنها تکیه نمودن به این روش بسیار خطرناک است و ممکن است باعث **publication bias** و انواع دیگر سوگیری‌ها شود.



## ۲-۴. جستجوی سایر مقالات مروری مرتبط

این شیوه نیز بسیار مفید است و کمک می‌کند که محقق با صرف زمان کم به تعداد قابل ملاحظه‌ای از منابع دست‌یابد. البته بایست دقت لازم به عمل آورده‌شود تا تنها از مقالات مروری باکیفیت بالا استفاده شود.

## ۲-۵. بانکهای اطلاعات کاغذی

همچنین این شیوه جستجو به محقق کمک می‌کند که به مقالات مرتبطی دست‌یابد که قبلاً از ایجاد بانکهای الکترونیک منتشر شده‌اند. مسلماً برای جستجوی منابع متعلق به دوره قبلاً از ایجاد بانکهای اطلاعات الکترونیکی تنها راه ممکن مراجعه به منابع کاغذی است.

Index Medicus از سال ۱۸۷۹ و Excerpt Medica از سال ۱۹۴۸ چاپ شده‌اند. همچنین Science Citation Index که فرم کاملتری از SciSearch میباشد از سال ۱۹۷۴ ایجاد و به چاپ رسیده است.

## ۲-۶. شناسایی منابع چاپ نشده

این قسمت سخت‌ترین نوع جستجو میباشد. چرا که اکثراً محقق هیچ روش تعریف شده‌ای برای تعیین استرانیژی جستجوی خود ندارد. یکی از روشهای پیشنهاد شده برای این قسمت از جستجو، مکاتبه نمودن با اشخاص سرشناس در حیطه مورد تحقیق و مراکز مربوطه میباشد. لیست محققین صاحب نام را میتوان از بین نویسندگان مقالات مرتبط استخراج نمود. همچنین مراکز تحقیقاتی را نیز میتوان با

یک جستجوی نظامند اینترنت به دست آورد. همچنین در حال حاضر بانکها اطلاعاتی تشکیل شده است که مشخصات کارآزمایی‌های بالینی در حال انجام را ثبت می‌کنند. از آنجایی که تعداد قابل توجهی از این تحقیقات به تولید هیچ مقاله‌ای منجر نمی‌شوند لذا جستجوی آنها می‌تواند بسیار با اهمیت باشد. بعضی از مهمترین این بانکها عبارتند از:

[www.controlled-trial.com](http://www.controlled-trial.com)

[www.trailsCentral.org](http://www.trailsCentral.org)

[www.clinicalTrial.gov](http://www.clinicalTrial.gov)

همچنین اطلاعات مربوط به عوارض داروها در سایتهای خاص همچون موارد زیر قابل دستیابی است:

[www.health.gov.au](http://www.health.gov.au)

[www.open.gov.uk/mca](http://www.open.gov.uk/mca)

البته سایتهای شرکتهای بزرگ دارویی نیز می‌توانند اطلاعات ارزشمندی در اختیار محققین قرار دهند. بجز موارد ذکر شده، گزارشات دولتی و سازمانهای بین‌المللی مانند سازمان بهداشت جهانی نیز ممکن است حاوی اطلاعات ارزشمندی باشند.

### ۲-۷. تدوین استراتژی جستجو

همانطور که قبلاً شرح داده شد مهمترین هدف جستجو بایست به حداکثر رساندن حساسیت **Sensitiviey** باشد ولی مسلماً به دلیل محدودیتهای زمانی، پولی و اطلاعاتی، امکان دستیابی به حساسیت ۱۰۰ درصد (یعنی گزینش ۱۰۰ درصد مقالات و نتایج تحقیقات مرتبط) تقریباً غیر ممکن است.

از طرف دیگر دستورالعمل جستجو برای منابع الکترونیک و دستی دقیقاً یکسال نیست. حتی این دستورالعمل برای بانکهای اطلاعاتی الکترونیک نیز تفاوت دارد ولی در کل میتوان یک الگوی کلی به شکل زیر ارائه نمود.

عبارت کامل جستجو معمولاً مرکب از سه جزاست که با ترکیب **and** به یکدیگر متصل میشوند. این سه جز عبارتند از:

الف- عبارت مربوط به بیماری مورد نظر مانند اسم علمی بدخیمی و یا ناراحتی مورد تحقیق. در بسیاری از موارد چندین اسم برای بیماری قابل تصور است و یا تظاهرات بالینی متعدد دارد. مثلاً نارسایی عروق کرونر میتواند با عبارتهای مختلف بیان شود. در این موارد کلیه اسامی قابل تصور ابتدا لیست شده و سپس با ترکیب **OR** به یکدیگر متصل میشوند. بدین شکل سعی میشود تا کلیه مقالاتی که به هر اسم بیماری مورد نظر را مطالعه نموده اند وارد مطالعه شوند.

ب- عبارت مورد نظر برای متغیر مستقل که عموماً همان تداخل انجام شده در کارآزماییهای بالینی هستند. این قسمت نیز همانند قسمت قبل ممکن است ترکیبی باشد و اسامی مختلف متغیر مستقل با ترکیب **OR** به یکدیگر متصل شوند.

ج- نوع تحقیق، به عنوان مثال اگر محقق در نظر دارد فقط مطالعات کارآزمایی بالینی را وارد نماید میتواند عبارت سوم را به گونه ای تنظیم کند که فقط این دسته از مطالعات انتخاب شوند.

البته به جز این سه قسمت اصلی بیان شده گاه نیاز است جمعیت هدف را نیز منطبق بر سوال پژوهش محدود نمود. در بعضی از موارد، سوال پژوهش در زیر گروه

خاص مانند خانمهای باردار یا کودکان است و یا از نظر سنی محدودیتهایی وجود دارد، لذا ممکن است محقق در استراتژی جستجوی خود این موارد را نیز مد نظر قرار دهد. از موارد بسیار مهمی که اکثراً توصیه میشود آن است که سال یا زبان مقالات به عنوان فاکتور محدود کننده از ابتدا در نظر گرفته نشوند چرا که تحقیقات متعدد نشان داده اند که این محدودیتهای میتوانند باعث سوگیری شوند. از موارد دیگر توصیه شده آن است که محققین سعی نمایند دو تیم جستجوگر مستقل تشکیل دهند و تیمها به صورت جداگانه منابع را جستجو نمایند تا بدین شکل مطمئن شوند که نتایج به اندازه کافی پایانی (Reliably) دارند. به صورت تجربی ثابت شده است که نتایج جستجوی افراد صاحب نظر با افراد کم تجربه متفاوت است و جالب آن است که در بعضی از موارد، افراد کم تجربه جستجوی بهتری انجام میدهند. این تفاوت شاید به دلیل اعمال نظرهای شخصی باشد که افراد با تجربه در کار به خرج میدهند.

بعد از جستجوی بانکهای اطلاعاتی و گزینش اولیه مقالات نوبت به آن میرسد که با مرور اجمالی خلاصه مقالات، تنها گروهی که مرتبط هستند حفظ شده و مقالات غیر مرتبط حذف شدند. البته در این مرحله معمولاً تعداد قابل ملاحظه ای از مقالات هستند که تصمیم گیری در خصوص آنها بر اساس خلاصه مقاله ممکن نیست. لذا توصیه میشود تا برای بالا بردن حساسیت جستجو، این دسته از مقالات نیز نگه داشته شده و تصمیم در خصوص ارتباط آنها با موضوع تحقیق بر اساس کل مقاله صورت گیرد.

در این مرحله نیز داشتن دو تیم مستقیم باعث ارتقا دقت خواهد شد. معمولاً تیمها به صورت مستقل، خلاصه مقالات را مطالعه و دسته بندی میکنند. در

مواردی که بین نظریات دو تیم اختلاف باشد گروه سوم نظر می‌دهد و تصمیم‌گیری از طریق بحث انجام خواهد شد. در این مرحله بعضی معتقدند که تیم‌های بررسی‌کننده باید مقالات را با کد بشناسند و اسم ژورنال و اسمی نویسندگان نباید در دسترس آنها قرار گیرد. (Blindness)، چراکه اسم ژورنال و حتی نویسنده در صورت مشهور بودن ممکن است تاثیرگذار باشد.

مراحل ذکر شده اگر چه می‌تواند به صورت کاغذی انجام شود ولی به دلیل زمان‌بر بودن و همچنین خطاهای انسانی توصیه می‌شود با استفاده از برنامه کامپیوتری همچون Procite, Reference manager و یا Endnote صورت پذیرد. این برنامه‌ها کمک می‌کنند تا به راحتی مقالات تکراری شناسایی و حذف شوند و کلیه مراحل انجام شده و نتایج آنها ذخیره شده و سرعت کار بالا رود.

البته برنامه‌های نوشته در نرم افزارهای MS-Access و Fox pro نیز وجود دارند که یک نمونه آن Meerkat بوده و قابلیت‌های اختصاصی بسیار خوبی برای انجام جستجوی ساختار یافته دارند. اطلاعات بیشتر در خصوص این نرم‌افزار در آدرس زیر قابل دسترس می‌باشد:

[www.update-software.com/meerkat](http://www.update-software.com/meerkat)

نرم افزار مذکور که با برنامه Access طراحی شده در حال حاضر توسط بسیاری از گروه‌های ککراین مورد استفاده قرار می‌گیرد.

## جستجوی ساختاریافته و متاآنالیز

جدول 2: تعداد مقالات یافت شده در جستجوی افراد متخصص در مدلاین و تعداد مقالات کارآزمایی های بالینی وارده شده در متاآنالیز

منبع	موضوع	تعداد مقالات گزینش شده در جستجوی اولیه	تعداد مقالات وارد شده در متاآنالیز
Dickerson(1985)	Inkerauenticadar hemorrhaye	36	19
	Hyperbiliubinemia	39	18
Benskein(1988)	Liver disease	9643	155
Ohlsson(1989)	pregnancy	125	10
Cotzsche(1991)	rheumatoidathrities	738	128
	Homeopothy	52	18
Kleijner(1992)	Vitamin c	81	22
	cingko	46	14
Dickerrin(1994)	ophthalmoloyy	1520	193
Lacy(unpubished)	newborn	21	8
Haghdoost(2004)	Interaction between <i>P.faciparum</i> and <i>P.vivax</i>	721	61

Moher و همکاران (۱۹۹۶) نشان دادند که کیفیت مقالات کار آزمایی بالینی چاپ شده به زبان انگلیسی و آلمانی در چند موضوع متاآنالیز تفاوت معنی داری وجود نداشته است.

### ۲-۸. سوگیری در انتشار (publication bias)

این اصطلاح برای نشان دادن احتمال بیشتر نوشتن، ارسال و همچنین شانس چاپ مقالات با نتایج معنی دار آماری است. بنابراین ممکن است ارزیابی مقالات چاپ

## جستجوی ساختاریافته و متاآنالیز

شده نتواند نمایانگر کل تحقیقات انجام شده در هر موضوع باشد. به عبارت دیگر این امکان وجود داشته- باشد که بدلیل چاپ نشدن مقالات با نتایج غیر معنی‌دار آماری، نتیجه‌گیری نهایی به سمت معنی‌دار جلوه دادن ارتباطات بین متغیرها باشد (جدول ۳).

**جدول ۳: وضعیت چاپ ۲۸۵ مطالعه ثبت شده در کمیته اخلاق مرکز تحقیقات اکسفورد بین ۱۹۸۷ الی ۱۹۸۴**

مطالعات با نتایج غیر معنی‌دار آماری <sup>۱</sup>		مطالعات با نتایج معنی‌دار آماری <sup>۱</sup>		وضعیت چاپ
درصد	تعداد	درصد	تعداد	
۳۴/۳	۴۵	۶۰/۴	۹۳	چاپ شده
۲۳/۷	۳۱	۲۴/۷	۳۸	ارائه شده در کنفرانس
۴۲	۵۵	۱۴/۹	۲۳	نه چاپ شده و نه ارائه شده
۱۰۰	۱۳۱	۱۰۰	۱۵۴	کل

۱. منظور  $p$ -value کمتر از ۰.۰۵ است.

منبع: Easterbrook et. al. (۱۹۹۱)

البته مدارکی در دسترس است که در نشان می‌دهد سالهای اخیر روند سوگیری انتشار در حال کاهش می‌باشد. البته تا زمانی که بتوان با اطمینان این خطا را قابل اغماض دانست احتمالاً سالها و حتی دهه‌ها فاصله است. بنابراین تا آن زمان برای حل این مشکل سه راه حل وجود دارد:

### الف- نادیده گرفتن سوگیری انتشار: این ساده-

ترین روش ممکن است که تعداد قابل ملاحظه‌ای از متاآنالیزهای چاپ شده از این راه حل استفاده نموده‌اند. با توجه به این نکته که این خطا قطعاً در بسیاری از جستجوهای ساختاریافته وجود

دارد لذا نادیده گرفتن آن شاید قابل توجیه نباشد.

البته در مواردی نیز می‌توان به آسانی ادله‌ای ارائه نمود که در موضوع مورد مطالعه این خطا قابل اغماض است. به عنوان مثال در متاآنالیز انجام شده توسط حقدوست (۲۰۰۴) برای بررسی تداخل بین سوش‌های مالاریا به نظر می‌رسد که احتمال چاپ نتایج تحقیقات ارتباطی به نتایج مطالعه ندارد. البته برای توجیه عدم وجود سوگیری در انتشار، روش‌های آماری نیز وجود دارد که در فصول بعدی به صورت مبسوط به آنها پرداخته خواهد شد.

**ب- تلاش برای یافتن اطلاعات غیر منتشر شده:** این روش اگرچه بهترین راه‌حل است ولی در عمل بسیار مشکل می‌باشد. چرا که یافتن سر نخ‌های مفید برای کشف مطالعات چاپ نشده در سطح دنیا بسیار دشوار است. البته امید است با روند سریع گسترش ارتباطات الکترونی، امکان تشکیل بانک‌های بین‌المللی برای ثبت عناوین تحقیقات و آدرس الکترونیک مراکز تحقیقات و محققین آسان شده و به این شکل امکان جستجوی آنها فراهم شود. ولی برای تحقق این آرزو نیز احتمالاً سالها زمان نیاز دارد.

**ج- استفاده از تکنیک‌های آماری برای شناسایی احتمال سوگیری در انتشار و کاهش اثر آن:** در حال حاضر می‌توان با کمک بعضی از انواع نمودارها و حتی آزمون‌های آماری احتمال وجود سوگیری در انتشار را تعیین و حتی سعی کرد تا حد امکان آنرا کاهش داد. با توجه به محدودیتهای دو روش دیگر این شیوه کاربرد وسیعی یافته است که در فصول بعدی به صورت مبسوط‌تر در



خصوص این تکنیکها بحث خواهد شد. تنها نکته قابل ذکر آن است که این روشها نیاز به بعضی پیش فرضها دارند که گاهی اوقات این مفروضات غیر قابل قبول میباشند. همچنین قدرت (Power) این آزمونها بستگی به تعداد مقالات دارد و در صورت کم بودن تعداد مطالعات وارد شده درآنالیز دقت پاسخ آنها زیر سوال خواهد رفت.

با توجه به این توضیحات باید دقت نمود که سوگیری انتشار یکی از محدودیتهای اصلی برای تعمیم نتایج مطالعات متاآنالیز به کل جامعه بوده و عمدتاً تاثیر آن زیادتیر جلوه دادن اثر فاکتورهای مورد مطالعه و ارتباطات بین متغیرها است.

## فصل سوم

### ۳. سنجش کیفیت مطالعات

---

شما در این فصل با این مفهوم آشنا می‌شوید که چرا سنجش کیفیت متدلوژیک مطالعات وارد شده در جستجوی ساختاریافته با اهمیت است و برای این ارزیابی باید به چه نکاتی توجه داشت. همچنین این فصل به شما کمک می‌کند تا به شیوه‌ای استاندارد، اطلاعات مورد نیاز خود را از گزارشات و یا مقالات استخراج نموده و برای متاآنالیز آماده نمایید.

بعد از جستجوی بانک‌های اطلاعاتی و منابع قابل دسترس و گزینش اولیه مقالات و گزارشات مورد قبول، نوبت به استخراج نتایج مطالعاتی است که در راستای سوال پژوهش هستند.

اطلاعات استخراج شده از مقالات به دو دسته قابل تقسیم هستند: دسته اول اطلاعاتی هستند که برای ارزیابی کیفیت مطالعات به کار می‌روند (quality assessment) و دسته دوم اطلاعاتی هستند که مستقیماً در راستای سوال پژوهش بوده و معمولاً میزان تاثیر فاکتور مورد مطالعه و میزان واریانس یا دامنه اطمینان آن را نشان می‌دهد.

هدف اصلی این قسمت از کار، استخراج اطلاعات با بالاترین پایایی (reliability) و اعتبار (validity) می‌باشد.

### ۱-۳. پایایی

پایایی به معنی میزان ثبات نتایج است و اختصاصاً در استخراج نتایج به معنی میزان ثبات نتایج ثبت شده برای متاآنالیز در میان کل نتایج ارائه شده است.

به صورت کلی می‌توان دو علت را برای کاهش پایایی نتایج استخراج شده بیان نمود. دسته اول خطاهای انسانی است؛ این خطا به خصوص در ارزیابی کیفیت مطالعات بیشتر مشهود است، چرا که بسیاری از این متغیرها بر اساس برداشت شخصی تعیین می‌شوند و به همین دلیل دیدگاه فردی و شرایطی که در آن مقالات مطالعه می‌شود می‌تواند در نتایج، تاثیر گذار باشد. دسته دوم خطاها مربوط به متن مورد مطالعه است. به عنوان مثال ممکن است بین نتایج مندرج در خلاصه مقاله و اصل نتایج، تفاوت‌هایی وجود داشته باشد و همچنین ممکن است بین توضیحات یک مقاله در خصوص روش کار با توضیحات مربوطه در گزارش مبسوط آن تفاوت وجود داشته باشد. بنابراین اهمیت دارد که مشخص شود منبع استخراج نتایج چه بوده است.

برای افزایش پایایی اطلاعات استخراج شده شاید بهترین پیشنهاد را بتوان در قالب موارد ارائه شده توسط Stock و همکاران (۱۹۸۲) خلاصه نمود که به شرح زیر می‌باشد:

۱. ایجاد فرمهای اولیه برای جمع‌آوری اطلاعات و استفاده آزمایشی از آنها قبل از شروع مطالعه اصلی. به عبارت دیگر بایست ابزاری قوی برای استخراج اطلاعات طراحی شود.

۲. تهیه کتابچه‌های راهنما با توضیحات مبسوط و ارائه مثال برای نحوه استفاده از فرمتها در استخراج اطلاعات.

۳. آموزش کامل و دقیق افرادی که مسئولیت استخراج اطلاعات را دارند برای استفاده از فرمها و نحوه استفاده از کتابچه‌ها.
۴. تعیین میزان همخوانی نتایج استخراج شده توسط افراد استخراج کننده اطلاعات به عنوان شاخصی از پایایی.
۵. اعمال هر گونه تغییری در فرمها و کتابچه‌های راهنما براساس اطلاعات بدست آمده از بند چهارم.
۶. تبیین راهکارهای اجرایی برای وارد نمودن افراد جدید برای استخراج اطلاعات.
۷. ترغیب افراد استخراج کننده اطلاعات به شرکت فعال در بحثها و تصمیم‌گیریهای مربوطه به ابزار جمع آوری اطلاعات و قوانین و دستورالعمل‌های مربوطه.

### ۲-۳. اعتبار

اعتبار در جستجوی ساختاریافته و متاآنالیز عبارت است از اینکه نتایج حاصل در مطالعات تا چه حد ممکن است بدلیل ایرادات متدلوژیک مطالعات انتخاب شده و یا شیوه اجرایی آنها از واقعیت دور شده باشند (Moher-۱۹۹۵). مسلماً نتایج مطالعات دقیقتر به حقیقت نزدیکتر است و لذا بایست در مرحله ارزیابی کیفیت مطالعات، نگاهی انتقادی به متدلوژی مطالعات وجود داشته باشد. البته در کنار این مورد بایست به نقش افراد استخراج کننده اطلاعات نیز توجه نمود. به عبارت دیگر از آنجایی که بعضی از متغیرهای سنجیده شده در مطالعات به صورت صد درصد عینی نیستند لذا نظر فرد استخراج کننده می‌تواند به صورت تاثیر گذار

نتایج مطالعات را تحت شعاع قرار دهد. به همین دلیل عده ای معتقدند که افراد استخراج کننده بایست اطلاعات مربوط به نام نویسندگان و ژورنال را نداشته باشد (blindness).

Berline و همکاران (۱۹۹۷) نشان دادند که چگونگی مخفی نمودن اسم نویسندگان و ژورنال در نتیجه‌گیری افراد استخراج کننده تاثیر می‌گذارد. البته علاوه بر این راهکار، روش های دیگری نیز توصیه می‌شود از جمله اینکه افراد ارزیابی کننده کیفیت متدلوژی مقالات با افراد استخراج کننده اطلاعات مربوط به متاآنالیز متفاوت باشند و یا حداقل افراد در دو زمان مختلف و با دو فرم متفاوت کیفیت متدلوژی مقالات و نتایج تحقیقات را استخراج نمایند (task separation).

اما به صورت کلی شاید بتوان چهار منبع اصلی خطا در متدلوژی مقالات را به شرح زیر عنوان نمود. البته این خطاها بیشتر با تاکید بر مطالعات کارآزمایی بالینی است که قابل تسری به سایر انواع مطالعات نیز می‌باشند.

۱. **سوگیری در گزینش نمونه‌ها (Selection bias):** این خطا در رابطه با مطالعات کارآزمایی بالینی زمانی بوجود می‌آید که نمونه‌ها و یا محقق قبل از ورود نمونه‌ها به گروه‌های درمانی مطالعه بتوانند در نحوه تقسیم آنها تاثیرگذار باشند. به عبارت دیگر افراد بعد از موافقت برای ورود به مطالعه بایست بصورت تصادفی در دو یا چند گروه مداخله (اصطلاحاً بازوهای مطالعه) وارد شوند. در غیر این صورت احتمال بروز این نوع سوگیری بسیار زیاد است.

برای حل این معضل راهکارهای اجرایی متعددی بیان شده است که علاقه مندان می‌توانند برای

مطالعه به کتابها و مقالات مربوط به نحوه تصادفی نمودن در کارآزمایی‌های بالینی مراجعه کنند.

این خطا در مطالعات کوهرت با حذف یا کاهش اثر فاکتورهای مخدوش‌کننده و در مطالعات مورد-شاهدی با انجام همسان سازی مورد و شاهد (Matching) قابل اصلاح است.

۲. **سوگیری در نحوه مراقبت از نمونه‌ها در گروه‌های مورد مطالعه (Performance bias):** معنی واضحتر این خطا در مطالعات کارآزمایی بالینی ارائه خدمات مراقبتی متفاوت در بازوهای مختلف درمانی است. این موضوع با مخفی نگه داشتن نوع تداخل در افراد شرکت‌کننده (Blindness) قابل کاهش است. در مطالعات کوهرت و مورد شاهدی نیز بایست دقت کافی در نحوه اندازه گیری متغیر مستقل (Exposure) صورت پذیرد.

۳. **سوگیری در ریزش نمونه‌ها (Attrition bias):** به معنی تفاوت قابل ملاحظه در حذف شدن نمونه‌ها در حین انجام مداخله و یا پیگیری (follow up) در گروه‌های مختلف مورد مطالعه است. در صورتی که نمونه‌های حذف شده (Missing data or loss to follow up) بدلیل بی‌اثر بودن درمان دریافت شده از مطالعات خارج شده- باشند، اطلاعات نمونه‌های باقیمانده مبین دقیق گروه‌ها نخواهند بود.

۴. **سوگیری در تشخیص نتیجه (Detection bias):** به معنی سوگیری در اندازه‌گیری متغیر وابسته است. به عنوان مثال اگر محقق در زمان ثبت نتایج درمانها بداند که هر فرد چه دارویی را دریافت- نموده ممکن است مقدار گزارش نموده اثر داروها را به شکلی که مورد نظر اوست تغییر دهد این

خطا را با مخفی نمودن نوع تداخل هر نمونه در مطالعات کارآزمایی بالینی و کوهرت میتوان کاهش داد. در مطالعات مورد-شاهدی نیز بایست دقت ویژه‌ای برای تعریف دقیق و استاندارد نمونه (Case definition) ارائه داد.

### ۳-۳. سنجش کیفیت مطالعات

همانگونه که بیان شد میتوان انتظار داشت که نتایج مطالعات دقیقتر به واقعیت نزدیکتر باشد، لذا بایست به شیوه مناسبی کیفیت مطالعات اندازه‌گیری شود. این سنجش از کار بسیار بااهمیت می‌باشد و مطالب بسیار زیادی در خصوص آن نگاشته شده است. در منابع مربوط به **evidence based medicine** این قسمت تحت عنوان **critical appraisal** آورده می‌شود. با این وجود متأسفانه بایست اقرار نمود که هنوز شیوه استاندارد و جامعی برای سنجش اعتبار نتایج مقالات وجود ندارد.

بصورت کلی میتوان روشهای موجود را به دو دسته تقسیم نمود:

۱. تقسیم نمودن مقالات به سه گروه: با احتمال کم، متوسط و زیاد سوگیری.

جدول 4 راهنمای ساده تقسیم‌بندی مقالات بر اساس میزان احتمال سوگیری در نتایج آنها		
خطر سوگیری	تفسیر	ارتباط با معیارهای ارزیابی مقالات
کم	احتمال انحراف نتایج بر اثر سوگیری کم است به دلیل احتمال انحراف	تقریباً تمام معیارها رعایت شده است
متوسط	نتایج بر اثر سوگیری بایست نتایج با احتیاط تفسیر شوند	حداقل یکی از معیارها تا حدودی رعایت شده است
زیاد	به احتمال زیاد نتایج بر اثر سوگیری از واقعیت منحرف شده اند	حداقل یکی از معیارها رعایت نشده است

این روش اگرچه ساده می‌باشد ولی در عمل مشکل آفرین است. چراکه معمولاً معیارهای ارزیابی کیفیت مقالات هم ارزش نیستند. همچنین در صورتی که تعداد معیارها زیاد شود استفاده از این روش دشوار می‌گردد.

۲. تدوین مقیاس عددی: مقیاسهای زیاد با دستورالعملهای متفاوت طراحی شده‌اند تا کیفیت مقالات را بصورت عددی تعیین نمایند. بعنوان مثال Moher(1995-6) برای سنجش کیفیت مطالعات کارآزمایی‌های بالینی ۲۵ مقیاس و ۹ چک لیست تدوین نموده‌اند. معمولاً این چک لیستها بین ۳ تا ۵۷ گزینه دارند و پر نمودن آنها بین ده دقیقه تا یک ساعت وقت می‌گیرند. همچنین هدف آنها نیز دقیقاً یکسان نیست. بعضی از این مقیاسها برای ارزیابی مقالات ارسالی برای چاپ در مجلات است در حالی که بعضی از آنها برای ارزیابی مقالات برای ورود به مطالعات مروری هستند.



نتیجه تقریباً تمامی این چک لیستها تعیین امتیاز عددی از کیفیت مقالات است و برای این کار امتیاز تک تک سوالات با یکدیگر جمع می‌شود. البته در بعضی از این مقیاسها به امتیازات سوالات قبل از جمع زدن وزن داده می‌شود. با این وجود (Jüni(1999 نشان داد که دقت بعضی از مقیاسهای پیچیده چندان بیشتر از مقیاسهای ساده نیست.

مشکل عمده این روشها آن است که بر اساس اطلاعات مندرج در مقاله یا گزارش، کیفیت مطالعات ارزیابی می‌شود و اگر موضوعی در آنها نوشته نشده باشد فرض بر عدم رعایت آن مورد گذاشته می‌شود. این نکته گاه باعث انحراف می‌شود. بدین معنی که ممکن است در اصل مطالعه به آن نکات توجه شده است ولی در مقاله و یا گزارشی بصورت کامل شرح داده نشده باشد. بنابراین متأسفانه بایست ابراز نمود که روش مناسبتر برای سنجش کیفیت اکثر مقالات وجود ندارد.

البته مشکل ارزیابی کیفیت مطالعات غیر کارآزمایی بالینی پیچیده تر است. چراکه در این مطالعات تصادفی نمودن گروهها برای یکسان سازی اثرات فاکتورهای مخدوش کننده (randomization) معنایی ندارد و از آنجایی که معمولاً تمامی فاکتورهای مخدوش کننده در مطالعات قابل شناسایی نیستند و یا اندازه گیری آنها سخت است، لذا با روشهای متدلوژیک و یا آماری امکان از بین بردن کامل اثرات مخدوش کنندگی در مطالعات نیمه تجربی و یا مشاهده ای (مورد-شاهدی و کوهرت) ممکن نیست و گاه این اثرات از مقدار ارتباط مشاهده شده بین متغیرها بیشتر است (Kunz-1995).

ولی به هر شکل نتایج اینگونه مطالعات در صورت صحیح انجام شده حاوی نکات ارزشمندی است که بایست در تصمیم گیریها مد نظر قرارگیرد. به همین دلیل تاکنون

معیارهای بسیار زیادی برای سنجش اعتبار مطالعات مشاهده‌ای طراحی شده است ( Horwitz 1979; Feinstein 1982; Levine 1994, and Bero 1999 ).

### ۳-۴. شیوه سنجش کیفیت مطالعات در عمل

معمولاً بحث اساسی در این قسمت آن است که چند نفر و چگونه، مقالات و گزارشات را مطالعه کنند و این افراد چه خصوصیات باید داشته باشند. معمولاً با افزایش افراد، دقت ارزیابی افزایش می‌یابد و احتمالاً پایایی بهتری بدست می‌آید. اما در عوض حجم کار زیاد شده و عدم همخوانی نظرات افراد (disagreement) می‌تواند مشکل آفرین شود. به همین دلیل عموماً توصیه می‌شود که هر مقاله و یا گزارش را حداقل دو نفر مطالعه نمایند. ولی قبل از آن بایست دقت نمود که افراد تیم کاملاً توجیه شده باشند و بعنوان تمرین، چند مقاله با کیفیت‌های خوب، متوسط و بد را بصورت گروهی مطالعه و ارزیابی نموده باشند. همچنین دستورالعمل شیوه ارزیابی، بایست کاملاً روشن تدوین شده و برای مواردی که اختلاف نظر بوجود می‌آید راه حل ارائه داده باشد.

مسلماً اعضا تیم باید با متدلوژی تحقیق آشنا بوده- باشند و در صورت لزوم دوره‌های آموزشی خاص را بگذرانند و از نظر علمی نیز با موضوع مورد بررسی بیگانه نباشند. ولی تحقیقات مختلف نشان داده است که ضرورت ندارد که ایشان از بین متخصصین و صاحب نظران آن رشته باشند.

در خصوص مخفی نمودن اسامی نویسندگان و اسم ژورنال نیز توافق کلی وجود ندارد. (Jadad 1996) نشان داد که مخفی نمودن اسامی اعتبار نتایج را افزایش می‌دهد در

حالی که Berlin (1997) نتیجه کاملاً متفاوتی بدست آورد و نشان داد که این کار تنها حجم کارهای اجرایی را افزایش می‌دهد. به هر شکل، بنظر می‌رسد بهترین توصیه آن باشد که محقق بر اساس شرایط و موضوع کار تصمیم بگیرد که آیا میزان خطای مذکور ارزش افزایش حجم کار و مخفی نمودن اسامی را دارد یا خیر.

مسئله مهم دیگر نحوه برخورد با امتیاز کیفیت مطالعات در نتیجه‌گیری نهایی است. بصورت کلی می‌توان بیان نمود که به چهار روش می‌توان با امتیازات مقالات برخورد نمود:

### ۱. تعیین امتیاز مشخصی برای ورود مقالات به

**مطالعه:** به عبارت دیگر تنها مقالاتی وارد مطالعه شوند که حداقل امتیاز ممکن را کسب نموده باشند.

### ۲. تقسیم‌بندی مقالات به زیر گروههایی بر اساس

**کیفیت آنها:** در بعضی مطالعات مروری و متاآنالیز، مقالات را بر اساس کیفیتشان به زیر گروههایی تقسیم می‌کنند و سپس در هر زیر گروه جداگانه آنالیزهای مربوطه را انجام می‌دهند (sub-group analysis).

### ۳. تعیین ارتباط بین امتیاز کیفیت مقالات و نتایج

**آنها:** به عبارت دیگر تعیین اینکه آیا نتایج گزارش شده در مقالات، ارتباطی با کیفیت آنها دارد یا خیر؟ برای پاسخ به این سوال معمولاً علاوه بر استفاده از نمودارهای ساده همچون نمودار پراکنش (scatter plot) از روشهای آماری نیز استفاده خواهد شد که در فصلهای بعدی شرح داده خواهد شد.

۴. وزن دادن مقالات بر اساس کیفیت آنها: در روش معمول در متاآنالیز وزن مقالات بر اساس واریانس شاخص مورد نظر محاسبه می‌شود. به عبارت دیگر مطالعاتی که شاخص مورد مطالعه را با دامنه اطمینان باریکتری تعیین نموده‌باشد وزن بیشتری می‌گیرد. در این روش بیان می‌گردد که معیار وزن دهی تنها واریانس نباشد بلکه امتیاز کیفیت نیز در تعیین وزن مطالعات دخالت داشته‌باشد.

هر یک از این روشها دارای امتیازات و محدودیتهای خاصی می‌باشند. در دو روش اول، تعداد مقالات وارد شده در آنالیز محدود شده و گاه بدلیل کم بودن مقالات امکان هر گونه کار آماری از بین می‌رود. در روش سوم نیز این سوال مطرح است که آیا انتظار رابطه خطی داشتن بین نتایج مقالات و کیفیت آنها فرض صحیحی است. در روش چهارم نیز مشکل اصلی یافتن روش آماری مناسبی برای ترکیب نمودن امتیاز کیفیت مقالات با واریانس شاخص مورد مطالعه است.

به هر شکل بسته به موضوع، تعداد مقالات، و کیفیت آنها، محقق بایست یک یا ترکیبی از روشهای فوق را بکار گیرد. البته در فصلهای بعدی به مباحث آماری آن بیشتر پرداخته خواهد شد.

### ۳-۵. رابطه بین احتمال انتشار مقاله و کیفیت مطالعه

شاید چندان دور از انتظار نباشد که مقالات با کیفیت بالاتر شانس چاپ بیشتری داشته‌باشند. معمولاً برای کارهای با کیفیت بالا هزینه بیشتری صرف می‌شود و تیم تحقیقاتی توان بیشتری در انجام پژوهش

## جستجوی ساختاریافته و متاآنالیز

دارد. لذا انتظار می‌رود که نتایج تحقیقات با کیفیت بهتر بیشتر منتشر شوند.

البته در این میان عوامل دیگری نیز نقش دارند. با توجه به تفاوت بسیار زیاد بین مجلات در قبول و چاپ مقالات، شاید چندان دور از انتظار نباشد که مقالات با کیفیت پایین‌تر نیز بتوانند برای خود مجوز انتشار بعضی از مجلات را اخذ نمایند. نمونه‌ای از این حالت را می‌توان در جدول زیر یافت.

**جدول 5** تعداد و درصدی از مطالعات کارآزمایی بالینی با کیفیتهای مختلف که توانسته‌اند بصورت کامل به چاپ برسند

تعداد مطالعه چاپ شده		تعداد مطالعه انجام شده	امتیاز کیفیت مطالعه
درصد	تعداد		
۳۰	۳	۱۰	کیفیت خوب
۳۷.۷	۴۳	۱۱۴	کیفیت متوسط
۳۴.۶	۱۸	۵۲	کیفیت پایین

Source: Chalmers et al. (1990)

این فرض را می‌توان از زاویه دیگری نیز بررسی نمود. همانگونه که در قسمت ۲-۸ شرح داده شد سوگیری در انتشار می‌تواند نتایج متاآنالیز را منحرف نماید. در آن قسمت بیان شد که معمولاً مطالعاتی که ارتباط آماری معنی‌داری را نشان می‌دهند شانس انتشار بیشتری دارند. اما در این قسمت بیان می‌شود که احتمال چاپ مقالات با کیفیت بالا بیشتر است. لذا در مواردی که مقالات با کیفیت پایین، هم‌زمان نتایج آماری معنی‌داری را نشان نداده باشند شانس بسیار کمتری برای چاپ می‌یابند. این فرضیه

را می‌توان بعد از جمع‌آوری اطلاعات و هنگام آنالیز  
نتایج بررسی نمود.

## فصل چهارم

### ۴. استخراج نتایج

---

این فصل در خصوص روش مناسبی برای استخراج نتایج مقالاتی است که از حداقل اعتبار برخوردار هستند و در آن نحوه تنظیم فرمهای مناسب استخراج نتایج شرح داده خواهد شد.

در تحقیقات رایج در حیطه علوم پزشکی، معمولاً نمونه انسانی، حیوانی و یا آزمایشگاهی واحد تحقیق را تشکیل می‌دهند و اطلاعات مربوط به تک تک واحدها از طریق مشاهده، مصاحبه، پرسشگری و یا استخراج از اطلاعات بایگانی شده جمع‌آوری می‌شود. به هر شکل حجم نمونه عبارت است از تعداد نمونه‌هایی که اطلاعاتشان جمع‌آوری می‌شود.

این مفهوم در مطالعات مروری ساختاریافته و متاآنالیز متفاوت است و گزارشات و مقالاتی که واجد شرایط هستند به عنوان نمونه وارد مطالعه شده و اطلاعاتشان استخراج می‌شود. به عبارت دیگر حجم نمونه در این گونه مطالعات برابر با تعداد مقاله‌ای است که گزینش شده است.

شیوه جمع‌آوری اطلاعات در مطالعات مروری ساختاریافته و متاآنالیز، عمدتاً از طریق استخراج اطلاعات نتایج مطالعه دیگران (گزارشات و مقالات) و ثبت در فرمهای مربوطه می‌باشد. اگرچه در بعضی موارد محقق مجبور است بعضی از اطلاعات را از طریق تماس با نویسندگان و حتی موسسات تحقیقاتی کسب و ثبت کند، چرا که ممکن است

همه اطلاعات لازم به صورت مکتوب در مقالات موجود نباشد.

با این توضیح روشن است که قبل از مرحله جمع‌آوری اطلاعات، محقق موظف است فرم مخصوص جمع‌آوری اطلاعات را تنظیم کند. این فرم بایست جامع بوده و کلیه اطلاعات ارزشمند و لازم برای پاسخ به اهداف تحقیق را پوشش دهد. اما از طرف دیگر بایست توجه داشت که جمع‌آوری اطلاعات اضافی باعث افزایش حجم کار و دشواری کار افراد استخراج‌کننده اطلاعات خواهد شد و از طرف دیگر ممکن است باعث افزایش حجم اطلاعات ناقص گردد چرا که ممکن است پاسخ تمامی سوالات در همه مقالات و گزارشها یافت نشود. حتی اطلاعات اضافه در مرحله تحلیل نتایج نیز دردسر آفرین بوده و باعث می‌شود که دقت و تمرکز در پاسخگویی به سوال پژوهش کم شود.

بر این اساس لازم است تا محقق در ابتدای کار و قبل از شروع مرحله جمع‌آوری اطلاعات، وقت و دقت کافی را برای تنظیم فرم مناسب جمع‌آوری اطلاعات مبذول دارد.

### ۴-۱. قسمتهای اصلی فرم جمع‌آوری اطلاعات

اگرچه بین فرمهای جمع‌آوری اطلاعات اختلافات اساسی وجود دارد ولی بعنوان یک راهنمای کلی می‌توان عنوان داشت که یک فرم خوب باید دارای قسمتهای زیر باشد.



## اطلاعات مربوط به مطالعه و فرد جمع‌آوری کننده اطلاعات

همانطور که قبلاً نیز شرح داده شد، ممکن است اطلاعات یک گزارش و یا مقاله توسط چند نفر و جداگانه استخراج شود. لذا بهتر است در فرم جمع‌آوری اطلاعات مشخص باشد که چه کسی اطلاعات مذکور را استخراج کرده است.

همچنین بهتر است روشن باشد که دقیقاً این اطلاعات از چه مقاله و گزارشی جمع‌آوری شده است. به همین دلیل ذکر نام مقاله و بعضی از مشخصات اصلی آن مانند اسم نویسنده اول، نام ژورنال و سال انتشار ثبت شود.

از آنجایی که ممکن است تمامی اطلاعات لازم در گزارشات و یا مقالات وجود نداشته باشد و برای بدست آوردن آنها از روشهای دیگر همچون تماس مستقیم با نویسنده و یا موسسه پژوهشی استفاده شده باشد، بایست جایی برای ثبت منبع اطلاعات نیز در نظر گرفته شود.

## اطلاعات مربوط به شرایط گزینش مطالعه

این دسته از اطلاعات شامل دو قسمت می‌باشند:

۱. اطلاعات مربوط به کیفیت مقاله، که در فصل قبل بصورت مبسوط شرح داده شدند. این اطلاعات باید در فرم جمع‌آوری اطلاعات همراه با تمامی معیارها آورده شوند. البته بعضی مواقع برای آسانتر شدن کار، اطلاعات مربوطه در فرم جداگانه‌ای ثبت می‌شود و تنها امتیاز کلی کیفیت مطالعه باشد در این فرم وارد می‌گردد.

۲. اطلاعات مربوط به مشخصات خاصی از مقاله که در گزینش دخالت داشته است. مثالهایی از این دسته می‌تواند نوع مطالعه، مدت پیگیری بیماران، نوع

مداخلات و درمانهای مورد مقایسه و گروه سنی نمونه‌ها باشد. ثبت این دسته از اطلاعات کمک می‌کند که در مراحل بعد و در صورت نیاز امکان تغییر بعضی از این شرایط ممکن باشد.

### مشخصات مطالعه

این دسته از اطلاعات شامل سولاتی در خصوص روش مطالعه، مشخصات نمونه‌های وارد شده در هر مطالعه، نوع تداخل انجام شده، و نوع و نحوه سنجش متغیر وابسته (outcome) می‌باشد. از آنجایی که متدلوژی همه مطالعات دقیقاً یکسان نیست و بعضی از این تفاوتها ممکن است توجیه‌کننده تفاوت‌های دیده‌شده در مطالعات مختلف باشند، ثبت این اطلاعات بسیار با اهمیت است.

گروه هدف مورد مطالعه (مانند گروه سنی، جنسی و نژادی) و همچنین نحوه شناسایی و حذف متغیرهای مخدوش‌کننده از مسایل مهمی هستند که می‌توانند بر نتایج تحقیق تاثیرگذار باشند.

اگر نوع مطالعه از جمله معیارهای ورود نباشد و محقق خود را محدود به نوع خاصی از مطالعات همچون کارآزمایی بالینی نداند، سایر اطلاعات مهم در این قسمت بر اساس نوع مطالعه متفاوت خواهد بود. بعنوان مثال نکات اساسی در مطالعه مورد-شاهدی نحوه گزینش نمونه‌ها و گروه کنترل است در حالی که در مطالعه کوهرت، نحوه دنبال نمودن نمونه‌ها و مدت مطالعه (follow up) با اهمیت است. بدین جهت فرم جمع‌آوری اطلاعات بایست این امکان را فراهم نماید که اطلاعات را متناسب با نوع مطالعه جمع‌آوری نماید.

اطلاعات مربوط به نوع مداخله نیز بسیار با اهمیت است. دز داروی تجویز شده، نحوه مصرف، مدت مصرف و

همراهی آن با سایر داروها در میزان تاثیر داروها بسیار موثر است. لذا بعضی از این تفاوتها ممکن است بصورت کاملاً موثری تفاوتهاي مشاهده شده در بین مطالعات را توجیه نماید.

متغیر وابسته اندازه‌گیری شده نیز از اهمیت زیادی برخوردار است. همانگونه که در فصل اول ذکر شد، گرفتاری قلبی-عروقی می‌تواند دامنه وسیعی داشته باشد و از گرفتاری‌های بدون علامت تا مرگ ناگهانی تغییر نماید و اگر هدف مطالعه تعیین اثر مصرف آسپرین در کاهش خطر حملات قلبی-عروقی است، متغیر وابسته یعنی حملات قلبی-عروقی ممکن است در مطالعات مختلف بسیار متفاوت باشد. به همین دلیل بایست محقق منطبق بر سوال پژوهش تصمیم بگیرد که تا چه حد تنوع در متغیر وابسته قابل قبول است و نوع آن را نیز دقیقاً ثبت نماید.

### ۲-۴. فرمهای الکترونیک و کاغذی

کاغذی و یا الکترونیک بودن فرم تاثیر به‌سزایی در روش جمع‌آوری اطلاعات دارد. به همین دلیل بهتر است با دقت مزایا و معایب هر یک از آنها را دانست.

فرمهای کاغذی آسانتر پر می‌شوند و استخراج کننده اطلاعات نیازی به آشنایی با کامپیوتر ندارد. همچنین در تمام شرایط و بدون امکانات اضافی و تنها با یک قلم و کاغذ قابل پر نمودن هستند. اما بایست توجه داشت که نهایتاً برای تجزیه و تحلیل باید اطلاعات وارد کامپیوتر شوند و در صورت استفاده از فرمهای کاغذی، یک مرحله ورود اطلاعات نیز بایست در نظر گرفته شود.

در مقابل فرمهای الکترونیکی، گاه پیچیده بوده و برای پر کردن آنها افراد باید آموزشهای بیشتری ببینند. همچنین فرد متخصص و آشنا به برنامه‌نویسی باید فرمها را طراحی و ارتباط آنها با بانکهای اطلاعاتی مربوطه را برقرار نماید.

اما بایست توجه داشت که بدلیل قابلیت‌های بیشتری که در فرمهای الکترونیکی است، سرعت پرکردن آنها بیشتر می‌باشد. بعنوان مثال می‌توان فرم را به گونه‌ای تنظیم نمود که با وارد نمودن يك كد، بصورت اتوماتيك مشخصات مقاله و با كدي ديگر مشخصات فرد پرکننده در فرم وارد شوند و یا با انتخاب نوع مطالعه بصورت اتوماتيك تنها سوالات مربوط به همان نوع مطالعه پرسیده شود.

یکی از موارد مهم دیگری که باید مد نظر باشد امکان مقایسه نمودن فرمهایی است که اطلاعات يك مقاله و یا گزارش را ثبت نموده‌اند. همانگونه که قبلاً نیز شرح داده شد برای بالا بردن دقت و پایایی نتایج بهتر است حداقل دو نفر بصورت مستقل هر مقاله را مطالعه و اطلاعات را استخراج نمایند. حال این سوال پیش می‌آید که کنترل فرمهای متفاوت کاغذی آسانتر است یا الکترونیکی؟

اگر تعداد فرمها محدود باشد معمولاً کنترل فرمهای کاغذی چندان دشوار نیست ولی با زیاد شدن فرمها و یا تعداد افراد ثبت‌کننده اطلاعات، کار با فرمهای کاغذی دشوارتر شده و خطاهای مربوطه بیشتر می‌شود. به همین دلیل استفاده از فرمهای الکترونیک ترجیح داده می‌شود.

### ۳-۴. انجام پیش مطالعه برای شناسایی مشکلات فرمها

مسلماً در تنظیم فرمها بایست از مشاوره افراد متخصص استفاده نمود و بر اساس اهداف مطالعه، فرمها را تنظیم نمود. در این تیم توصیه می‌شود حداقل یک نفر آشنا با مراحل بعدی کار یعنی تجزیه و تحلیل نتایج و مباحث آماری وجود داشته‌باشد. با این وجود نمی‌توان مطمئن بود که فرم اولیه تهیه شده جوابگویی تمام نیازها باشد و در عمل ممکن است نقایص آن روشنتر شود. به همین دلیل انجام یک پیش مطالعه (pilot study) بسیار مفید است. در این قسمت پیشنهاد می‌شود که محقق اصلی و یکی از افراد تیم جداگانه اطلاعات چند مقاله را استخراج و در فرمهای تهیه شده وارد نمایند و سعی نمایند در آن مرحله ایرادات فرمها را شناسایی کنند. در مورد تعداد نمونه‌های لازم برای پیش مطالعه نمی‌توان عدد دقیقی بیان کرد ولی مسلماً هرچه حجم آن بیشتر باشد بهتر است و معمولاً بصورت تجربی حجم نمونه آن را برابر ۵ تا ۱۰٪ نمونه نهایی در نظر می‌گیرند.

از موارد مهم دیگر کنترل فرمهای الکترونیکی است. چراکه متأسفانه در بسیاری از موارد این فرمها دارای ایراداتی است که در عمل باعث ایجاد دردها عدیدهای می‌شوند. لذا بهتر است با حوصله، نحوه کار کردن برنامه و فرمها در حالت‌های مختلف کنترل شوند.

## فصل پنجم

### ۵. آماده‌سازی اطلاعات برای متاآنالیز

در این فصل توضیح داده خواهد شد که چگونه بایست اطلاعات استخراج شده از مطالعات اولیه را یک دست نموده و برای متاآنالیز آماده نمود. همچنین سعی می‌شود به صورت مختصر به نکات آماری و همچنین متدلوژیک مهم در تفسیر نتایج مطالعات اولیه اشاره گردد.

همانگونه که بیان شد در تحقیقات رایج علوم پزشکی، نمونه مورد مطالعه عمدتاً انسان، بافت انسانی و یا حیوان آزمایشگاهی است. ولی در متاآنالیز، نتایج هر مطالعه به عنوان یک نمونه محسوب می‌شود. به عبارت دیگر می‌توان متاآنالیز را آنالیز تجمعی نتایج مطالعات دیگر دانست. به همین دلیل مطالعات بر روی انسانها، بافتهای انسانی و یا حیوانات آزمایشگاهی را مطالعات اولیه (primary) و متاآنالیز را گاه مطالعات ثانویه (secondary) می‌نامند. البته در خصوص کلمه "مطالعات ثانویه" اختلاف نظر وجود دارد و بعضی این کلمه را برای مطالعاتی استفاده می‌کنند که اطلاعات یک مطالعه اولیه را مجدد و با یک تکنیک جدید آماری تحلیل می‌کند و یا به دنبال روابط جدیدی در اطلاعات گذشته می‌گردد. به دلیل این اختلاف نظر، در این کتاب از کلمه متاآنالیز بجای لفظ "مطالعات ثانویه" استفاده خواهد شد.

در مرحله آماده‌سازی اطلاعات مطالعات اولیه برای متاآنالیز، بایست بر روی اطلاعات استخراج شده،

عملیاتی صورت پذیرد. اگرچه در متاآنالیز سعی می‌شود مقالات مشابه گزینش شوند، ولی حتی در صورت یکسان بودن متدولوژی مطالعات، ممکن است شیوه تجزیه و تحلیل نتایج و نحوه ارایه آنها، یکسان نباشد و یا شاخصهای متفاوتی را برای ارایه نتایج خود برگزیده-باشند.

به همین دلیل در مرحله آماده‌سازی، سعی می‌شود اطلاعات استخراج شده تا حد امکان یکسان شده و امکان ترکیب آنها با استفاده از شیوه‌های آماری متاآنالیز فراهم گردد.

برای انجام متاآنالیز، محقق بایست شاخص یکسانی برای خلاصه نمودن نتایج کلیه مطالعات اولیه تولید نماید. بعضی از مطالعات اولیه ممکن است نتایج خود را به شیوه یکسان و با شاخص مشترکی گزارش نموده‌باشند ولی در اکثر اوقات محقق باید خود یک شاخص را تعیین و برای همه مطالعات اولیه محاسبه نماید.

علاوه بر آن، بایست میزان انحراف معیار ( و یا انحراف استاندارد) شاخص محاسبه شده نیز تعیین گردد. انحراف معیار مذکور برای وزن دادن نتایج مطالعات در متاآنالیز لازم می‌باشد.

در ادامه این فصل سعی می‌شود مهمترین و پر مصرفترین شاخصهای مورد استفاده در متاآنالیز شرح داده‌شده و کاربردهای آنها بیان گردد.

### ۵-۱. شاخص و پارامتر

پارامتر (Parameter) به مقدار یک متغیر در کل جامعه گفته می‌شود در حالی که شاخص (Index) به مقدار آن متغیر در نمونه مورد بررسی اطلاق می‌گردد. بعنوان مثال میانگین وزن هنگام تولد نوزادان ایرانی، یک

پارامتر محسوب شود و مقدار محاسبه شده این میانگین در نمونه نوزادان مورد بررسی، شاخص می‌باشد.

هدف اصلی تحقیقات، تعیین پارامترها است، ولی در عمل بدلیل عدم امکان اندازه‌گیری پارامترها (از طریق مطالعه کل جامعه)، شاخصها محاسبه می‌شوند و از روی آنها، پارامترها تخمین زده می‌شوند. برای این هدف، یک نمونه تصادفی از جامعه گرفته شده و مقدار متغیر مورد نظر در نمونه‌ها اندازه‌گیری و توصیف می‌شود. سپس با استفاده از روشهای برآورد آماری (statistical estimation) مقدار پارامتر تخمین زده می‌شود.

### ۵-۲. شاخصهای قابل اندازه‌گیری در مطالعات توصیفی

در مطالعات توصیفی، هدف صرفاً برآورد یک پارامتر ساده است. بصورت کلی می‌توان بیان داشت که رایج‌ترین پارامترهای مورد سنجش در مطالعات توصیفی دو دسته هستند:

۱. میانگین: برای توصیف متغیرهای عددی (کمی) همچون وزن، قد، فشارخون و قند خون بکار می‌رود. به عبارت ساده‌تر، برای توصیف متغیرهای عددی معمولاً از میانگین استفاده می‌شود و به میانگین محاسبه شده در نمونه، شاخص و میانگین واقعی آن متغیر در کل جامعه (که معمولاً نامشخص است)، پارامتر گفته می‌شود.

۲. فراوانی یا درصد: برای توصیف متغیرهای کیفی (چند حالته و حتی بعضی متغیرهای رتبه‌ای) از فراوانی و یا درصد استفاده می‌شود، مانند فراوانی پرفشاری خون، فراوانی کمردرد، فراوانی ناهنجاری مادرزادی در نوزادان و فراوانی گروه



خونی AB. با کمی اغماض می‌توان تمام تحقیقات اپیدمیولوژیکی را که به دنبال تعیین شیوع و بروز بیماریها هستند را در این گروه قرار داد.

### ۳-۵. شاخصهای قابل محاسبه در مطالعات تحلیلی

در مطالعات تحلیلی، محقق یا در صدد نشان دادن ارتباط بین متغیرها است (رابطه فشارخون با متغیرهایی مانند سن، سطح خونی کلسترول، مصرف سیگار) و یا سعی می‌نماید تا اختلافات بین گروهی را نشان دهد (مانند اختلاف فشارخون مردان و زنان).

در تفسیر نتایج مطالعات تحلیلی باید دو مفهوم را کاملاً درک نمود که عبارتند از: معنی‌داری آماری (statistical significance) و شدت اثر (effect size or clinical importance difference).

### معنی‌داری آماری

معنی‌داری آماری نشان می‌دهد که آیا ارتباط و یا اختلاف مشاهده شده بین متغیرها تنها با شانس و تصادف قابل توجیه است یا خیر. بعنوان مثال اگر میانگین فشارخون سیستولیک اندازه‌گیری شده در نمونه‌های مرد و زن به ترتیب ۱۲۱ و ۱۱۸ میلیمتر جیوه باشد، این سوال می‌تواند مطرح شود که آیا مشاهده ۳ میلیمتر جیوه تفاوت در بین فشارخون سیستولیک مردان و زنان می‌تواند مبین تفاوت واقعی در بین کل مردان و زنان جامعه مورد بررسی باشد؟ به زبان دیگر آیا این امکان وجود دارد که در کل جامعه، میانگین فشارخون سیستولیک مردان و زنان دقیقاً یکسان بوده باشد و صرفاً

بدلیل شانس و تصادف در نمونه‌های وارد مطالعه شده، تفاوت ۳ میلیمتری دیده‌شده‌باشد؟ این احتمال توسط  $p$ -value نشان داده می‌شود و معنی آن این است که تفاوتها و یا ارتباطات مشاهده‌شده تا چه میزان می‌تواند صرفاً بر اثر شانس و تصادف بوده‌باشد؛ لذا کوچک بودن  $p$ -value نشان دهنده اثر کمتر تصادف بوده و تلویحاً بیان می‌کند که ارتباطات و یا اختلافات مشاهده‌شده در بین متغیرها ذاتی است و بر اساس شانس نیست. به این مفهوم در آمار، اصطلاحاً معنی‌داری (significance) گفته می‌شود.

در مثال فوق اگر  $p$ -value، ۰.۰۱ گردد نشان می‌دهد که در صورت فرض نمودن تساوی فشارخون سیستولیک کل مردان و زنان جامعه، فقط یک درصد امکان مشاهده تفاوت ۳ میلیمتر جیوه‌ای بین فشارخون سیستولیک نمونه‌های مردان و زنان صرفاً بدلیل شانس و تصادف است. از آنجایی که این احتمال کم است، می‌توان نتیجه گرفت که فشار خون مردان و زنان در جامعه مساوی نیست.

### شدت اثر

بجز معنی‌داری آماری، لازم است که میزان اختلافات و یا شدت ارتباطات بین متغیرها نیز مد نظر باشد. بعنوان مثال، تحقیقات متعدد نشان داده‌اند که سیگار کشیدن خطر سرطان ریه و معده را افزایش می‌دهد و این ارتباطات از نظر آماری معنی‌دار است. با این وجود، بر اساس مقدار  $p$ -value به راحتی نمی‌توان تعیین نمود که سیگار کشیدن چه میزان احتمال بروز این سرطانها را افزایش می‌دهد و کدام یک از این دو سرطان با شدت بیشتری با سیگار کشیدن مرتبط است.

برای بیان شدت ارتباطات و اختلافات، شاخصهای دیگری وجود دارد که در ادامه به مهمترین آنها اشاره خواهد شد. این شاخصها بصورت کمی شدت ارتباطات و اختلافات را نشان میدهند و این امکان را فراهم میآورند تا خواننده به راحتی روابط بین متغیرها را درک نماید.

البته ذکر این نکته لازم است که بین معنیداری آماری و شدت اثر ارتباط وجود دارد. در صورت زیاد شدن شدت اثر، مقدار  $p$ -value کاهش مییابد و نشان میدهد که احتمالاً روابط مشاهده شده بر اثر شانس و تصادف نیست. البته بایست مد نظر داشت که در این بین، حجم نمونه نیز نقش بسیار مهمی را ایفا میکند. بعنوان مثال در بررسی "اثر زندگی در ارتفاعات بر خطر بروز گرفتگی عروق کرونر"، در یک نمونه کوچک مثلاً ۲۰ نفری، صرفاً بدلیل شانس و تصادف ممکن است در افرادی که در ارتفاع زندگی میکنند این خطر حدود دو برابر دیده شود و به عبارتی شدت اثر ۲ از نظر آماری معنیدار نشود. در مقابل با یک حجم نمونه بالا مانند ۲۰۰ نفر، شدت اثر حتی کمتر از ۲ ممکن است از نظر آماری معنیدار شود.

لذا میتوان بیان نمود که در مطالعات تحلیلی، صرف دیده شدن یک ارتباط معنیدار آماری کافی نیست. بلکه بایست میزان ارتباط نیز تعیین شود. در صورتی که این ارتباط بسیار قوی باشد، با حجم کم نمونه نیز میتوان  $p$ -value معنیدار یافت ولی در صورتی که شدت اثر کم باشد بایست حجم نمونه بیشتری داشت تا به سطح معنی-داری دست یافت.

آزمونهای ساده مورد استفاده در اکثر مقالات مانند آزمون کای-دو،  $t$ -test، ANOVA، تنها میتوانند معنیداری ارتباطات و اختلافات را نشان دهند و برای تعیین شدت اثر بایست از روشهای کاملتر اپیدمیولوژیک و آماری

استفاده نمود به همین دلیل مفهوم شدت اثر کمتر در ذهن محققین نقش گرفته است. اما با توجه به اهمیت این مفهوم و تاثیر آن بر تصمیم‌گیری‌ها و همچنین در متاآنالیز، در فصل بصورت مشروح در این خصوص بحث خواهد شد.

### شاخصهای نشان دهنده شدت اثر

شاخصهای مذکور بر اساس نوع متغیرهای مورد بررسی متفاوت هستند. یعنی بایست بر اساس نوع متغیر تعیین نمود که بهترین شاخص نشان دهنده ارتباط بین متغیرها چه می‌باشد. در ادامه به صورت مختصر این شاخصها، جداگانه شرح داده می‌شوند.

### متغیرهای دو حالته (dichotomous)

در بسیاری از مطالعات تحلیلی، هدف بررسی ارتباط بین بیماریها و عوامل خطری است که دو حالته هستند، بعنوان مثال رابطه بین خطر بروز بدخیمی و مواجهه با بعضی مواد شیمیایی. در این گونه مطالعات هر دو متغیر وابسته و مستقل دو حالته هستند. حتی در صورت دو حالته نبودن متغیر مستقل، بسیاری از محققین ترجیح می‌دهند که آن را بصورت دو حالته در آورده و تجزیه و تحلیل نمایند. مثلاً در مطالعه بین رابطه مصرف سیگار با بدخیمی‌ها، با وجود آنکه مقدار مصرف سیگار به خودی خود متغیر عددی است، ولی در بسیاری از مطالعات بر اساس تعاریفی مشخص، آن را به متغیر دو حالته (سیگاری/غیر سیگاری) و یا چند حالته (سیگاری/ غیر سیگاری/ ترک کرده) تبدیل می‌نمایند.

شاخصهای مورد استفاده برای نشان دادن ارتباط بین متغیرهای دو حالت را میتوان برای نشان دادن رابطه یک متغیر دو حالت (مانند وجود یا عدم وجود بیماری) با یک متغیر اسمی چند حالت (مانند گروه خونی)، یا با یک متغیر رتبه‌ای (مانند سطح سواد) و حتی با یک متغیر عددی (مانند سن) نیز استفاده کرد. همچنین با استفاده از روشهای پیشرفته‌تر، حتی میتوان رابطه یک متغیر رتبه‌ای را با سایر انواع متغیرها با شاخصهایی که در ادامه این قسمت آورده خواهند شد سنجید.

بصورت کلی برای نشان دادن ارتباط بین متغیرهای دو حالت چندین شاخص آماری و اپیدمیولوژیک وجود دارد که مهمترین آنها عبارت هستند از:

### خطر نسبی (RR) - (Relative Risk or Risk Ratio)

این شاخص نشان میدهد که خطر بروز بیماری (یا نتیجه مورد مطالعه) در گروه مواجه‌یافته با عامل خطر چند برابر خطر بروز در گروه مواجه نیافته است. به عنوان مثال اگر RR بین مصرف سیگار و بدخیمی معده ۲ باشد یعنی آنکه در گروه سیگاری خطر بروز این بیماری دو برابر گروه غیر سیگاری است.

برای درک ساده‌تر این شاخص و نشان دادن نحوه محاسبه آن میتوان از جدول زیر کمک گرفت:

کل	بلی	خیر	بیماری عامل خطر
A+B	B	A	منفی
C+D	D	C	مثبت

در کتابه آئی RR دمیولوژی به ادومفه وم جداگان ه اسد تفاده م شی ومعم دولأنت ون آمریک ایی RR مع ادل Relative Risk در متون انگلیس یی مع ادل Risk Ratio ار م یی رودر بعض یی م وارد RR مع ادل Rate Ratio نیز بکار برده شده است در متن م ورد مطالعه ه، دقت ش ودک ه RR ادل دقیر ق ک دام عی ارت بکار برده شده است.

A+B+C+D	B+D	A+C	کل
---------	-----	-----	----

در جدول فوق هر خانه نشان دهنده تعداد افرادی است که در آن گروه قرار می‌گیرند. به عنوان مثال D نشان دهنده تعداد بیمارانی است که با عامل خطر مذکور مواجه داشته‌اند، و A نشان دهنده افراد سالمی است که با عامل خطر مورد نظر برخورد نداشته‌اند.

بر این اساس خطر بروز بیماری در گروه‌هایی که با عامل خطر برخورد داشته‌اند و نداشته‌اند به ترتیب برابر است با:

$$\text{Risk of disease in exposed group} = \frac{D}{D+C}$$

$$\text{Risk of disease in Non - exposed group} = \frac{B}{A+B}$$

و نهایتاً مقدار خطر نسبی برابر خواهد بود با:

$$RR = \frac{\text{risk of disease in exposed group}}{\text{risk of disease in Non - exposed group}} = \frac{\frac{D}{D+C}}{\frac{B}{B+A}}$$

مقدار RR بیشتر از یک، نشان دهنده رابطه مثبت بین عامل خطر مورد نظر و بیماری است و در صورتی که RR کمتر از یک باشد رابطه بین عامل خطر و بیماری منفی است و به عبارت دیگر مواجهه با عامل مورد نظر، خطر بروز بیماری را کاهش می‌دهد. بر این اساس اگر RR برابر یک باشد نشان می‌دهد که عامل خطر مورد نظر رابطه‌ای با بیماری ندارد.

برای بررسی معنی‌داری RR محاسبه شده، می‌توان دامنه اطمینان ۹۵٪ آن را حساب نمود، اگر در دامنه مورد نظر عدد یک قرار گیرد نشان می‌دهد که رابطه بین عامل خطر و بیماری معنی‌دار نیست. بعنوان مثال اگر RR بین مصرف سبزی و بدخیمی روده برابر ۰.۷۵ و دامنه اطمینان ۰.۵۷-۰.۹ شود نشان می‌دهد که مصرف سبزی خطر بروز این بدخیمی را حدود ۲۵٪ کاهش می‌دهد و دامنه اطمینان آن یک را در بر نمی‌گیرد. به عبارت دیگر اگر مصرف سبزی خطر را ۲۵٪ کاهش ندهد به احتمال ۹۵٪ حداقل ۱۰٪ و حداکثر ۴۳٪ کاهش می‌دهد و از نظر آماری معنی‌دار است. حال اگر RR بین مصرف قرص آهن در دوران بارداری و خطر بروز زردی نوزادی برابر ۱.۰۲ با دامنه اطمینان ۰.۹ تا ۱.۴ شود، چون در این دامنه عدد یک قرار می‌گیرد، می‌توان نتیجه گرفت که رابطه بین این دو از نظر آماری معنی‌دار نیست.

#### تفاوت خطر (Risk Difference-RD)

این شاخص تفاوت بین خطر بروز بیماری را در گروهی که با عامل خطر مورد نظر مواجه داشته‌اند با گروه غیر مواجه نشان می‌دهد. به عبارت دیگر این مقدار برابر است:

$$RD =$$

*Risk of disease in exposed Group – Risk of disease in Non exposed group*

$$= \frac{D}{D+C} - \frac{B}{B+A}$$

بر اساس مثبت یا منفی بودن RD می‌توان در خصوص رابطه مثبت یا منفی بین بیماری و عامل خطر نظر داد. همچنین اگر دامنه اطمینان این شاخص عدد صفر را در بر گیرد، نشان می‌دهد که بین آنها رابطه آماری معنی‌داری وجود ندارد.

### نسبت شانس (Odds Ratio-OR)

این شاخص در ابتدا بیشتر در شرطبندیها مورد استفاده قرار می‌گرفت ولی به تدریج بدلیل خصوصیات ممتاز آمایش، مورد توجه زیاد متخصصین آمار حیاتی و اپیدمیولوژی قرار گرفت.

Odds یا شانس بنا به تعریف عبارت است از نسبت احتمال موفقیت به احتمال عدم موفقیت. به عنوان مثال

شانس عدد شش آمدن یک تاس برابر  $\frac{1}{5}$  است، چرا که

احتمال عدد شش آمدن  $\frac{1}{6}$  و احتمال عدد غیر شش آمدن  $\frac{5}{6}$

است و نسبت این دو مقدار برابر  $\frac{1}{5}$  خواهد بود.

حال اگر شانس بروز بیماری در گروه مواجهه با عامل خطر را بر شانس بروز بیماری در گروه عدم مواجهه با عامل خطر تقسیم نماییم مقدار حاصل برابر OR خواهد شد. به عبارت دیگر OR برابر است با

$$OR = \frac{\text{Odds of disease in exposed group}}{\text{Odds of disease in Non - exposed group}} = \frac{\frac{D}{C}}{\frac{B}{A}} = \frac{AD}{BC}$$



اگرچه OR از نظر مقدار و مفهوم با RR متفاوت است ولی تفسیر مقدار OR بسیار شبیه RR است. یعنی مقدار کمتر از یک، نشان دهنده اثر محافظتی و مقدار بیشتر از یک نشان دهنده اثر بیماری‌زایی عامل خطر است. OR برابر یک نشان دهنده عدم ارتباط عامل خطر با بیماری است و اگر دامنه اطمینان محاسبه شده برای OR عدد یک را در بر گیرد نشان دهنده عدم ارتباط معنی‌دار بین آنها است.

از نظر مقدار نیز OR در مقایسه با RR همیشه از یک فاصله بیشتری دارد. یعنی در صورت ارتباط مثبت بین عامل خطر و بیماری، مقدار عددی OR بیشتر از RR است و در صورت ارتباط منفی، مقدار OR کمتر از RR خواهد بود. هرچه شیوع بیماری مورد نظر بیشتر شود این تفاوت نیز بیشتر خواهد شد.

ولی از طرف دیگر دامنه اطمینان OR وسیعتر از RR می‌باشد و این تفاوت دامنه‌های اطمینان، برای شیوع بالای بیماری مورد نظر، بیشتر است.

اگر نوع مطالعه اجازه دهد، به راحتی و با استفاده از فرمولهای زیر می‌توان این دو شاخص را به یکدیگر تبدیل نمود.

$$RR = \frac{OR}{1 - (BR \cdot (1 - OR))} \quad \text{and} \quad OR = \frac{RR(1 - BR)}{1 - (BR \cdot RR)}$$

*BR = Baseline Risk or Risk in Non – expoused group*

یکی از مزایای OR در مقایسه با RR آن است که OR خاصیت جابجایی دارد یعنی اگر جای عامل خطر و بیماری جابجا شود مقدار عددی OR تغییر نمی‌یابد، در حالیکه RR این خصوصیت را ندارد و به راحتی نمی‌توان مقدار RR را با جابجا نمودن متغیر وابسته و مستقل تعیین

کرد. به عنوان مثال اگر OR مصرف مواد مخدر به عنوان عامل خطر برای سطح اقتصادی پایین، برابر ۳.۵ باشد، مقدار OR سطح اقتصادی پایین به عنوان عامل خطر مصرف مواد مخدر نیز ۳.۵ خواهد بود.

### متغیرهای عددی

در بسیاری از موارد متغیرهای وابسته، مستقل و یا هر دو از نوع عددی می‌باشند. بسته به نوع این متغیرها ممکن است مطالعات اولیه اطلاعات خود را به شیوه‌های مختلف آماری تجزیه و تحلیل نموده و گزارش نمایند. در ادامه به مهمترین روشها اشاره می‌شود.

### ضریب همبستگی

رابطه متغیرهای عددی را معمولاً با ضریب همبستگی می‌سنجند. در صورت وجود شرایط لازم، بهترین ضریب همبستگی ضریب پیرسن است که به صورت اختصاری با  $r$  نمایش داده می‌شود. از نظر عددی مقدار  $r$  بین -۱ تا +۱ تغییر می‌کند. عدد صفر نشان دهنده عدم وجود ارتباط است و علامت مثبت نشان دهنده ارتباط مثبت بین دو متغیر مورد نظر است. مقدار این شاخص را می‌توان بر اساس فرمول زیر محاسبه نمود:

$$r = \frac{\sum xy}{\sqrt{\sum (x^2)(y^2)}} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \cdot \sum (y - \bar{y})^2}}$$

در فرمول فوق مقدار  $x$  و  $y$  مقادیر متغیرهای مورد بررسی در هر نمونه می‌باشد.

در بعضی مطالعات ممکن است یک متغیر را در قبل و بعد از یک مداخله اندازه‌گیری نموده باشند و بجای ضریب

همبستگی بین مقادیر قبل و بعد از مداخله، مقدار میانگین و انحراف معیار آن را گزارش نموده باشند. در این گونه موارد و برای انجام متآنالیز بایست مقدار ضریب همبستگی را بر اساس اطلاعات اعلام شده در مقاله حساب نمود. این محاسبات را می‌توان با استفاده از فرمولهای زیر انجام داد:

$$r = \frac{SD(B)^2 + SD(A)^2 - SD(C)^2}{2 \cdot SD(B) \cdot SD(A)}$$

در فرمول فوق  $SD(A)$ ,  $SD(B)$  و  $SD(C)$  برابر با مقدار انحراف معیار متغیر مورد نظر قبل از مداخله (baseline)، بعد از مداخله (after) و انحراف معیار تغییرات (changes) است.

### رابطه بین متغیر دو حالت و عددی با استفاده از تفاوت میانگینها

هدف بسیاری از مطالعات و کارآزمایی‌های بالینی مقایسه یک متغیر عددی در دو (حتی چند گروه) است. بعنوان مثال برای بررسی اثربخشی یک داروی جدید در کاهش فشارخون، در کارآزمایی‌های بالینی، میزان تاثیر داروی مذکور با داروی استاندارد مقایسه می‌شود. میزان تفاوت فشارخون دو گروه بعد از دریافت داروها، می‌تواند نشان دهد که میزان اثر بخشی داروی جدید در مقایسه با داروی استاندارد چگونه بوده است.

برای متآنالیز اطلاعات اینگونه مطالعات، میانگین و انحراف معیار متغیر عددی (در مثال فوق فشارخون) در گروه‌های مختلف لازم است. معمولاً اکثر مقالات این شاخصها را بیان می‌کنند. ولی در صورت عدم گزارش انحراف معیار، می‌توان آن را بر اساس سایر شاخصهای گزارش شده محاسبه نمود. به عنوان مثال می‌توان از

فرمولهای زیر استفاده نمود و با استفاده از انحراف استاندارد ( $Se$ )، دامنه اطمینان، و یا فاصله بین صدک ۱۲۵ام و ۱۷۵ام (به شرط متقارن بودن توزیع متغیر) می-توان مقدار انحراف معیار را محاسبه نمود.

$$SD = Se \cdot \sqrt{n}$$

$$SD = \sqrt{n} \cdot \frac{(UL - LL)}{3.92}$$

$$SD = \frac{IQR}{1.35}$$

UL and LL are Upper and Lower Limits of 95% Confidence Interval

IQR is Inter Quartile Range

همچنین در بعضی مطالعات ممکن است مقدار  $p$ -value را گزارش نموده باشند که می-توان مقدار عدد  $t$  و یا  $z$  را از جداول مربوطه استخراج و سپس  $SD$  را محاسبه نمود.

شاخص تفاوت میانگینها به راحتی حساب می-شود و با کم کردن میانگین متغیر مورد مطالعه در دو گروه بدست می-آید. البته می-توان به جای محاسبه میانگین ساده از میانگین وزنداده شده نیز استفاده نمود. با این تعریف تفاوت میانگینها برابر است با:

*Mean Difference =*

$\bar{x}_i - \bar{x}_c$  where  $\bar{x}_i$  and  $\bar{x}_c$  are means intervention and control groups

رابطه بین متغیر دو حالت و عددی با استفاده از تفاوت استاندارد شده میانگینها (SMD)

برای محاسبه این شاخص، تفاوت میانگینهای محاسبه شده بر انحراف معیار تفاوتها تقسیم می-شود و از نظر

مفهوم، می‌توان آن را شبیه  $z$ -score تفسیر نمود که در آمار، کاربردهای فراوان دارد.

تفاوت استاندارد شده میانگینها این امکان را فراهم می‌کند تا تفاوت‌های ایجاد شده در متغیرهای مختلف را مورد مقایسه قرار داد. بعنوان مثال اگر بخواهیم میزان تاثیر يك دارو را در کاهش LDL, HDL، کلسترول و تري‌گلیسرید بصورت همزمان بسنجیم، طبیعتاً تفاوت ساده میانگینها ارزشمند نیست و می‌توان با استاندارد نمودن این تفاوتها، مقایسه‌های مذکور را انجام داد.

گاهی اوقات بعضی مقالات ممکن است يك متغیر عددی را بصورت دو حالت در آورده و OR را محاسبه نموده- باشند و بعضی از مقالات تفاوتها را با استفاده از SMD گزارش نموده باشند. برای یکسان سازی نتایج این گونه مقالات برای متاآنالیز می‌توان از فرمولهای زیر استفاده نمود که توسط Chinn (2000) شرح داده شده‌اند:

$$SMD = \frac{\sqrt{3}}{\pi} \log OR$$

$$Se(SMD) = \frac{\sqrt{3}}{\pi} Se(\log OR)$$

### تعداد وقوع يك پدیده

در بعضی موارد متغیر وابسته، تعداد به وقوع پیوستن يك حادثه و یا اتفاق است نه صرفاً وجود یا عدم وجود آن. به عنوان مثال ممکن است هدف يك مطالعه بررسی اثر يك داروی جدید در تعداد حملات تشنج باشد. در این حالت بعضی از نمونه‌ها ممکن است بیش از يك حمله را تجربه نمایند. به دلیل همبستگی‌های درون نمونه‌ای (within subject correlation) تجزیه و تحلیل این نتایج به

راحتی موارد قبلی نیست و نمی‌توان به راحتی از OR و RR استفاده نمود. بهترین شاخصها برای توصیف و تحلیل اینگونه نتایج، میزان نسبی (Relative Rate or Rate Ratio) و یا تفاوت میزانها (Rate Difference) است.

بنا به تعریف، میزان (Rate) عبارت است از تعداد حوادث مشاهده شده تقسیم بر شخص-زمان مورد بررسی. حاصل تقسیم میزان بروز یک پدیده در دو گروه برابر با میزان نسبی و اختلاف آنها برابر با تفاوت میزانها می‌باشد.

بعنوان مثال اگر ۱۵ حمله تشنج در ۱۰۰ نفر-ماه پیگیری بیماران گروه مصرف کننده یک داروی جدید و ۳۰ حمله تشنج در ۱۰۰ نفر-ماه پیگیری بیماران گروه مصرف کننده داروی استاندارد ثبت شده باشد، می‌توان گفت که میزان نسبی تشنج در گروه مصرف کننده داروی جدید نسبت به داروی استاندارد ۰.۵ است.

تفسیر این شاخص از بسیاری جهات بسیار شبیه RR می‌باشد که در قسمتهای قبل شرح داده شده است.

بر اساس فرمول (Whiteheel 1991) انحراف استاندارد مورد نیاز برای متاآنالیز میزان نسبی را می‌توان بر اساس این فرمول محاسبه نمود:

$$Se(\log \text{ rate ratio}) = \sqrt{\frac{1}{A} + \frac{1}{C}}$$

در این فرمول A و B برابر است با تعداد حوادث مشاهده شده در دو گروه مورد مقایسه.

همچنین (Hasselblad 1995) فرمول زیر را برای محاسبه انحراف استاندارد تفاوت میزانها بیان نموده است

$$Se(\text{rate difference}) = \sqrt{\frac{A}{x^2 + y^2}}$$

در این فرمول منظور از  $x$  و  $y$  ، نفر-زمان پیگیری نمونه‌ها در دو گروه می‌باشد.

### زمان تا وقوع يك پدیده

گاهی افراد مورد مطالعه، تا زمان وقوع يك حادثه پیگیری می‌شوند و مدت زمان مذکور ثبت می‌شود. در این گونه مطالعات، مدت پیگیری افراد اکثراً یکسان نیست و بعضی از افراد ممکن است به مدت بسیار طولانی و بعضی به مدت کمتر در مطالعه، وارد شده باشند. همچنین ممکن است افراد مورد مطالعه، به دلایل مختلف از مطالعه خارج شده باشند. در این گونه مطالعات، بهترین شیوه تجزیه و تحلیل نتایج استفاده از روشهای تحلیل بقا (survival analysis) می‌باشد. نسبت مخاطره (Hazard Ratio-HR) مهمترین شاخصی که می‌تواند تفاوت بین میزان خطر بروز يك پدیده را در دو یا چند گروه با لحاظ کردن زمان وقوع آن حادثه را نشان دهد. تفسیر این شاخص بسیار شبیه میزان نسبی است و روش آماری Cox model مقدار آن را همراه با انحراف استاندارد محاسبه می‌کند. در مواردی که از روش logRank Analysis برای تجزیه و تحلیل استفاده شده باشد، می‌توان با استفاده از فرمولهای زیر مقدار HR و انحراف استاندارد آن را محاسبه نمود:

$$HR = \frac{\text{Number of Observed events} - \text{Number of expect events}}{\text{variance}}$$

$$Se(HR) = \frac{1}{\sqrt{\text{variance}}}$$

#### ۴-۵. مقیاس شاخصها در متاآنالیز

تعدادی از شاخصهای شرح داده شده در قسمتهای قبل نسبی بوده و محاسبه آنها بر اساس تقسیم دو شاخص توصیفی میباشند. مهمترین شاخصها در این دسته،  $\text{Ratio Odds}$ ،  $\text{Risk Ratio}$  و  $\text{Hazard Ratio}$  و  $\text{Rate Ratio}$  میباشند.

دامنه اطمینان این دسته از شاخصها متقارن نیست. به عبارت دیگر فاصله بین مقدار محاسبه شده شاخص با دامنه بالا و پایین مساوی نیست. بعنوان مثال اگر  $\text{OR}$  برابر ۱.۴ محاسبه شده باشد، دامنه اطمینان آن مثلاً بین ۱.۰۲ و ۱.۷۵ است. به این دلیل و البته به دلیل بعضی از خصوصیات ذاتی این شاخصها، برای متاآنالیز بایست ابتدا لگاریتم طبیعی ( $\text{natural logarithm}$ ) آنها محاسبه شده و تجزیه و تحلیل بر روی آنها صورت گیرد (شرح بیشتر در فصلهای بعد آورده شده است).

این تغییرات در مورد سایر شاخصها لازم نیست و مستقیماً میتوان آنها را به همراه انحراف معیارشان در متاآنالیز وارد نمود.

#### ۵-۵. شاخصهای خام و تصحیح شده

برای سنجش روابط بین متغیرها میتوان از شاخصهای شرح داده شده استفاده نمود. ولی در محاسبه این شاخصها در مطالعات اولیه، ممکن است با اثر متغیرهای مخدوش کننده یکسان برخورد نشده باشد.



در خصوص نحوه تاثیر فاکتورهای مخدوش کننده بر نتایج مطالعات اولیه و خطاهای حاصل مطالب بسیار زیادی نوشته شده است ولی این اثرات در نتایج مطالعات متاآنالیزی چندان شرح داده نشده است. ولی آنچه مسلم است آن است که ترکیب شاخصهای خام و تصحیح شده صحیح نبوده و باعث زیر سوال رفتن نتایج مطالعات متاآنالیز می شود.

برای حل این مشکل بعضی از متخصصین تاکید دارند که در صورت مشهود بودن اثر متغیرهای مخدوش کننده چه به لحاظ آماری و چه به لحاظ بیولوژیک، نتایج خام به هیچ وجه نباید در متاآنالیز وارد شود. شاید یکی از دلایلی که اکثر متاآنالیز بر روی نایج کارآزمایی های بالینی صورت می گیرد نیز همین موضوع باشد چراکه انتظار می رود با اعمال صحیح روشهای تصادفی نمودن، اثرات متغیرهای مخدوش کننده به شدت کاهش و حتی حذف شود.

به هر شکل بایست دقت نمود تا حد امکان از وارد نمودن نتایج خام و یا ترکیب نتایج مقالاتی که نتایج خام و تصحیح شده را ارائه نموده اند اجتناب نمود.

البته روشهای نیز وجود دارد که بر آن اساس می توان تا حدودی و بدون داشتن ریز اطلاعات مقالات اولیه، نتایج خام را به نتایج تصحیح شده تبدیل نمود، ولی با توجه به محدودیتهای این روشها و پیش فرضهای آنها، کمتر از این روشها در مقالات متاآنالیز استفاده شده است.

آخرین نکته آن است که حتی نتایج مطالعات کارآزمایی های بالینی مشابه نیز گاه بدلیل اثر نمونه های خارج شده از مطالعه به راحتی قابل ترکیب نمودن نیستند. چراکه روشهای متفاوت آماری برای کاهش سوگیری حاصل از حذف نمونه ها در تحلیل نتایج این مطالعات وجود

دارد که مهمترین آنها روشهای (ITT) Intention to treatment و مدل سازی برای پیشبینی اطلاعات حذف شده است.

بدین لحاظ برای ترکیب نمودن نتایج اینگونه مطالعات بایست از نظر متدلوژی و آماری مطالعات به دقت بررسی شوند.

## ۵-۶. آخرین حربه برای متاآنالیز مطالعات غیر متجانس

با توجه به مطالب فوق، گاه به هیچ شکل امکان یکسان نمودن و آماده سازی نتایج مطالعات اولیه برای متاآنالیز وجود ندارد و بدلیل تفاوتی پایه ای بیشتر محقق به توصیف نتایج مطالعات و دسته بندی آنها می پردازد.

در این گونه موارد صرفاً به جهت بیان آماری جهت ارتباط و بدون امکان کمی نمودن شدت ارتباط، می توان از روش Vote Counting استفاده نمود که توسط (Cooper 1980) and Antman (1992) شرح داده شده است. در این روش، تعداد مقالاتی که رابطی مثبت در بین متغیرهای مورد نظر نشان داده اند را بر تعداد مقالاتی که رابطه منفی نشان داده اند تقسیم می کنند. سپس به وسیله آزمون غیر پارامتریک علامت (sign test) معنی دار بودن این نسبت را می سنجند.

در فرمهای پیچیده تر این روش، به جای تقسیم ساده تعداد مقالات مثبت و منفی، تعداد وزن داده شده آنها را بر اساس حجم نمونه، میزان دقت آماری و یا شدت اثر گزارش شده تعیین می کنند.

## فصل ششم

---

### ۶. مفاهیم اولیه متاآنالیز

---

این فصل ابتدا به بعضی از خصوصیات متاآنالیز می‌پردازد و کاربردهای آن را شرح داده و محدودیتهای آن را بر می‌شمارد. سپس به مبانی آماری متاآنالیز پرداخته و به زبان ساده به تشریح مراحل انجام متاآنالیز از زاویه آماری می‌پردازد. در این فصل سعی می‌شود مفاهیم پایه‌ای همانند مدل‌های فیکس و تصادفی و روش‌های استاندارد متاآنالیز به تفصیل بیان شوند.

اگرچه اکثراً متاآنالیز بر روی نتایج مطالعات تحلیلی صورت می‌گیرد، ولی این شیوه می‌تواند نتایج مطالعات توصیفی را نیز روی هم ریخته و یک نتیجه کلی ارائه دهد. در ادامه ابتدا به کاربردهای متاآنالیز در مطالعات توصیفی و تحلیلی اشاره می‌شود.

#### ۶-۱. متاآنالیز در مطالعات توصیفی

در مطالعات توصیفی سعی می‌شود که میانگین و یا فراوانی یک متغیر در جامعه اندازه‌گیری شود. بعنوان مثال ممکن است هدف تعیین میانگین فشارخون در یک منطقه باشد و یا هدف تعیین فراوانی پرفشاری خون باشد. مطالعات اولیه می‌توانند دامنه اطمینان شاخص مورد نظر را نیز تعیین کنند، مثلاً اگر هدف مطالعه تعیین شیوع پرفشاری خون در شهر کرمان باشد، و از

بین ۱۰۰۰ نمونه گرفته شده، ۱۶۰ نفر مبتلا دیده شود، نهایتاً می‌توان گفت که فراوانی پرفشاری خون در این جامعه ۱۶٪ درصد بوده و دامنه اطمینان ۹۵٪، این فراوانی از ۱۳.۷٪ الی ۱۸.۳٪ خواهد بود. این دامنه اطمینان تلویحاً نشان می‌دهد که شیوع پرفشاری خون در این منطقه اگر دقیقاً ۱۶٪ نباشد به احتمال ۹۵٪ عددی بین ۱۳.۷٪ الی ۱۸.۳٪ خواهد بود.

حال اگر اطلاعات مشابه‌ای از فراوانی پرفشاری خون در سایر نقاط کشور در دسترس باشد، می‌توان این نتایج را با یکدیگر ترکیب نموده و در قالب یک متاآنالیز تعیین نمود که بهترین تخمین از پرفشاری خون در سطح کشور چقدر است و همچنین آیا تفاوت‌های دیده شده در مطالعات مختلف در حد قابل قبول<sup>۳</sup> است و یا نشان دهنده آن است که این شاخص در سطح کشور یکسان نیست. همچنین در صورت مشاهده تفاوت، می‌توان بررسی نمود که آیا این تفاوت به دلیل فاکتورهای شناخته شده‌ای که در مطالعات اندازه‌گیری شده‌اند می‌باشد یا دلیل دیگری دارد.

با توجه به مثال فوق، می‌توان بیان نمود که در متاآنالیز مطالعات توصیفی که بر روی شاخص‌های متناظر استخراج شده از مطالعات متفاوت صورت می‌گیرد، سه هدف عمده دنبال می‌شود که عبارت هستند از:

۱. تخمین بهترین عدد از شاخص مورد نظر بعد از ترکیب نتایج مطالعات اولیه
۲. تعیین میزان توافق (consistency) نتایج مطالعات اولیه
۳. تعیین فاکتورهای احتمالی ایجاد کننده اختلاف بین نتایج مطالعات اولیه

<sup>۳</sup> منظور از حد قابل قبول، حدی است که تنها بر اساس شانس و تصادف قابل توجیه است.

## ۶-۲. متاآنالیز در مطالعات تحلیلی

هدف از متاآنالیز نتایج مطالعات تحلیلی با متاآنالیز مطالعات توصیفی متفاوت است و در آن سعی می‌شود که رابطه بین متغیرها مورد بررسی و تحلیل قرار گیرد.

### اهداف متاآنالیز مطالعات تحلیلی

برای روشنتر شدن این بخش، به این مثال توجه کنید. در خصوص اثر بخشی واکسن BCG در کاهش خطر ابتلا به سل تحقیقات مختلفی صورت گرفته که نتایج آنها بسیار متفاوت می‌باشد. در سال ۱۹۹۴، Colditz et al. متاآنالیزی بر روی نتایج مطالعات مداخله‌ای منتشر شده انجام دادند. این متاآنالیز نشان داد که احتمالاً تلقیح واکسن، خطر را حدود ۵۰٪ کاهش می‌دهد. ولی شدت اثر مشاهده شده در بین ۱۳ مطالعه انتخاب شده، بسیار متفاوت بود. به عبارتی بعضی مطالعات این اثربخشی را زیاد و بعضی کم مشاهده نموده بودند. در بررسی علت دیده شدن این تفاوت، محققین دریافتند که بین شدت اثر واکسن و عرض جغرافیایی رابطه معنی‌دار آماری وجود دارد و با افزایش عرض جغرافیایی میزان اثر واکسن نیز افزایش می‌یابد و سپس سعی نمودند که این رابطه را با فراوانی میزان آلودگی با میکوباکتریوم‌های آتیپیک توجیه نمایند.

بر اساس این مثال، می‌توان بیان داشت که متاآنالیز مطالعات تحلیلی سعی می‌کنند تا به پنج سوال اصلی پاسخ دهند که این سوالات عبارتند هستند از:

۱. جهت و میزان اثر چگونه است؟
۲. آیا اثر مذکور از نظر آماری معنی‌دار است؟

۳. آیا شدت اثر در بین مطالعات مختلف همخوانی دارد؟

۴. در صورت عدم همخوانی شدت اثر در بین مطالعات مختلف، آیا می‌توان دلیلی برای توجیه تفاوتها یافت؟

۵. آیا مستندات کافی برای قبول اثر مذکور وجود دارد؟

البته بایست توجه داشت که پاسخ به سوال آخر تنها از جنبه آماری مقدور نیست و بایست به سایر جنبه‌ها از جمله کیفیت مطالعات و تاثیر فاکتورهای مخدوش‌کننده نیز توجه داشت.

### گسترده‌گی رابطه مورد بررسی

از نظر گسترده‌گی رابطه بین متغیرها، در مطالعات متاآنالیزی ممکن است یک رابطه ساده و مشخص را مورد ارزیابی قرار دهند و یا ممکن است بسیار گسترده بوده و متغیرهای مستقل و وابسته ترکیبی داشته‌باشند. بر این اساس و از این زاویه نیز بصورت کلی می‌توان چهار نوع مشخص متاآنالیز را شرح زیر توصیف نمود:

۱. اکثر مطالعات متاآنالیزی رابطه بین دو متغیر مستقل و وابسته مشخص را مورد ارزیابی قرار می‌دهند و یا اثربخشی دو درمان مشخص را مقایسه می‌کنند. بعنوان مثال به دنبال آن هستند که رابطه پر فشاری خون را بر خطر مرگ در اثر نارسایی قلبی بسنجند و یا میزان تاثیر دو درمان دارویی مشخص را در کنترل فشارخون مورد مقایسه قرار دهند. در این مثالها، تعریف متغیر مستقل و وابسته کاملاً مشخص و محدود می‌باشد.

۲. در بعضی متاآنالیزها بیش از چند نوع درمان را مورد مقایسه قرار می‌دهند و هدف آنها تعیین بهترین نوع درمان است. بعنوان مثال به دنبال آن هستند که بهترین درمان موضعی را برای قارچهای انگشتان و ناخنهای پا تعیین نمایند. برای پاسخ به این سوال، عمدتاً محقق باید تمام ترکیبهای دوجه دو ممکن را با یکدیگر مقایسه نماید. بعنوان مثال اگر سه درمان رایج وجود داشته باشد، محقق بایست ۳ مقایسه را انجام دهد (مقایسه درمانهای ۱ با ۲، ۲ با ۳ و ۱ با ۳). از آنجایی که با زیاد شدن تعداد درمانها، تعداد مقایسه‌ها به شدت افزایش می‌یابد و همچنین امکان یافتن مقالات مربوطه کاهش می‌یابد، در عمل نتایج اینگونه متاآنالیز ضعیفتر می‌شود.

۳. در دسته‌ای دیگر، ممکن است تعریف متغیرهای مستقل و وابسته بسیار وسیع باشد. مثلاً هدف بررسی نقش برنامه‌های آموزش تلویزیونی بر اختلالات رفتاری کودکان باشد. در این مثال، متغیر مستقل، برنامه‌های تلویزیونی است که خود می‌تواند حالت‌های بسیار زیادی به خود بگیرد و از برنامه‌های کوتاه تا بلند، و همچنین آموزشهای مستقیم تا آموزشهای غیر مستقیم تغییر یابد. از طرف دیگر اختلالات رفتاری نیز بسیار گسترده بوده و شامل تعداد زیادی از مشکلات می‌شود. ترکیب نمودن همه این حالتها در یک متاآنالیز بسیار سوال برانگیز بوده و ممکن است به راحتی باعث انحراف خواننده شود. به همین دلیل در این گونه مطالعات، محقق بایست به صورت مستند توجیه نماید که چرا و به چه علت گزینه‌های مختلف را با یکدیگر ترکیب نموده و نتیجه حاصل چگونه بایست تفسیر شود.

۴. هدف بعضی متاآنالیزها، بررسی رابطه بین شدت اثر و بعضی متغیرهای مطالعات اولیه میباشد. بعنوان مثال محقق به دنبال آن است که دز مناسب داروهای استروئیدی استنشاقی را در درمان آسم تعیین نماید و به همین دلیل اطلاعات مطالعات اولیه را متاآنالیز مینماید. البته باید بیان نمود که تعداد محدودی متاآنالیز با هدف اصلی مذکور تدوین میشوند ولی در بسیار از متاآنالیزها، در کنار هدف اصلی، نگاهی مشابه به اطلاعات میشود.

### ۳-۶. نقش متاآنالیز

به واقع کارآیی یک متاآنالیز و نقش آن در پاسخ به سوال تحقیق بستگی به عوامل متعددی دارد که در فصول قبل شرح داده شد. ولی به صورت کلی میتوان چهار نقش اصلی برای متاآنالیز قائل شد:

۱. **افزایش توان آماری (statistical power):** بصورت مختصر، توان آماری عبارت است از احتمال کشف یک رابطه آماری معنی‌دار در صورت وجود آن رابطه. توان آماری رابطه مستقیمی با حجم نمونه دارد و با افزایش حجم نمونه احتمال کشف رابطه بیشتر میشود. به عبارت دیگر در صورت پایین بودن حجم نمونه ممکن است مطالعه نتواند رابطه بین متغیرها را کشف نماید. بدین لحاظ ممکن است هیچ یک از مطالعات اولیه به اندازه کافی حجم نمونه نداشته باشند و به همین دلیل توان آماری آنها زیر سوال باشد. ولی در متاآنالیز، چون نتیجه-گیری آماری بر اساس ترکیب نتایج مطالعات اولیه است، توان آماری افزایش می‌یابد.



۲. **افزایش دقت:** چون دقت تخمین میزان اثر نیز با حجم نمونه رابطه دارد، لذا دامنه اطمینان حاصل از مطالعات متاآنالیز همیشه دقیقتر<sup>۴</sup> از دامنه اطمینان مطالعات اولیه است.

۳. **پاسخ به سوالاتی که مطالعات اولیه نمیتوانند به آنها پاسخ دهند:** در مطالعات اولیه ممکن است بعضی متغیرها مانند گروه سنی و یا جنسی نمونه-ها، و یا دوز داروی مصرفی ثابت باشند. به همین دلیل هر یک از مطالعات اولیه به تنهایی نمی-توانند به تاثیر متغیرهایی که ثابت بوده اند پاسخ دهند، این در حالی است که در متاآنالیز بدلیل تفاوت‌های بین مطالعات اولیه، می-توان به نقش و تاثیر آنها پرداخت. نمونه اینگونه ارتباطات، در مثال قبل در خصوص اثر عرض جغرافیایی بر میزان اثر بخشی واکسن BCG آورده-شد.

۴. **ارزیابی دقیقتر تفاوت‌های بین یافته‌های مطالعات:** وجود پاسخ‌های متفاوت و حتی متناقض بین یافته‌های مطالعات مختلف امر چندان ناشایعی نیست. بسیار اتفاق می‌افتد که مطالعات کاملاً مشابه به یک سوال پژوهش متفاوت پاسخ می‌دهند. برای جمع‌بندی بهتر و دقیقتر نتایج مطالعات مختلف، و حتی بررسی آماری دلایل تفاوت‌های بین مطالعات، انجام متاآنالیز یکی از کارآمدترین روشها است.

<sup>۴</sup> منظور از دقت در دامنه اطمینان، فاصله بین حد پایین یا بالا دامنه با تخمین نقطه‌ای است. به عنوان مثال اگر ر شیوع پرفش آری‌زدون ۵٪ (بین نقطه‌ای پوشش آری‌زدون) اطمینان ۹۵٪ بدین ۱۲٪ تا ۱۸۱٪ تعیین شود، می‌توان بیان نمود که دقت تخمین ۳٪ است چرا که می‌توان دامنه اطمینان را به ۳٪±۱۵٪ نیز نوشت.

## ۴-۶. موارد نابجا استفاده از متاآنالیز

اگرچه انجام متاآنالیز می‌تواند در بسیاری از موارد، پاسخهای دقیقی به سوالات پژوهشگران بدهد ولی دارای نقاط ضعفی نیز می‌باشد که عدم توجه به آنها می‌تواند منجر به نتیجه‌گیری غلط شود. از جمله مهمترین این موارد، می‌توان به حالتی اشاره نمود که محقق نتایج تحقیقاتی که با یکدیگر تفاوت بسیار دارند را ترکیب نموده است. ترکیب مطالعات متنوع اولیه باعث می‌شود که محقق نتواند سوال پژوهش خود را به درستی شکل دهد و لذا پاسخهای ارائه شده نیز دقیق نخواهند بود. بعنوان مثال اگر محقق بخواهد عوامل خطر همه بدخیمی‌های پوستی را در یک متاآنالیز بررسی نماید، بدلیل تنوع انواع بدخیمی‌های و همچنین داشتن عوامل خطر متفاوت، به راحتی نمی‌تواند همه آنها را در یک ظرف ریخت و بصورت کلی عوامل خطر همه را یک جا لیست نمود. به عبارت دیگر نمی‌توان سیب و پرتقال را با یکدیگر مخلوط نمود و با هم وزن کرد.

از موارد دیگری که گاه در متاآنالیز مشکل آفرین می‌شود، تنوع روشهای اجرا و تجزیه و تحلیل نتایج مطالعات اولیه است. به عنوان مثال بعضی مطالعات ممکن است سعی نموده باشند تا در روش اجرا و یا تجزیه و تحلیل آماری اثر بعضی عوامل مخدوش کننده را حذف کنند؛ در مقابل بعضی از مطالعات نتوانسته باشند این عوامل را شناسایی و اثرشان را حذف نمایند. بدیهی است که ترکیب نتایج این مطالعات خالی از اشکال نیست و بایست به گونه‌ای مستدل در ترکیب نتایج آنها دقت نمود.

از جمله مشکلات دیگر متاآنالیز، تاثیر کیفیت پایین بعضی مطالعات اولیه بر کل نتایج است. اگر یک یا چند

مطالعه در مقایسه با بقیه از نظر متدلوژی تحقیق، دقت کافی را نداشته باشند باعث می‌شود تا نتیجه کل متاآنالیز زیر سوال رود.

و نهایتاً، سوگیری در انتشار و یا حذف تعدادی از مطالعات بدلیل زبان مقاله می‌تواند نتایج را مخدوش کند. این ایراد در فصول قبل به تفصیل بحث شده است.

### ۵-۶. مقدمات آماری متاآنالیز

به صورت ساده می‌توان گفت که متاآنالیز سعی می‌کند تا نتایج مطالعات اولیه را ترکیب نماید. به عبارت دیگر برآیند نتایج تحقیقات اولیه مختلف را نشان دهد. به عنوان مثال فرض کنید که ۱۰ مطالعه اولیه، رابطه بین میزان مصرف سبزی بر بروز بدخیمی معده را سنجیده اند. فرض کنید که کمترین و بیشترین اثر محافظتی مشاهده مصرف سبزی در این مطالعات بین ۱۰٪ تا ۶۰٪ باشد. حال سوال این است که برآیند نتایج این مطالعات چگونه بایست محاسبه شود؟

احتمالاً ساده‌ترین پاسخ به این سوال، میانگین درصدهای مذکور در ۱۰ مطالعه اولیه است. اگر پاسخ شما این باشد بایست گفت که شما مفهوم اصلی متاآنالیز را درک نموده‌اید!

حال بگذارید کمی دقیقتر موضوع را دنبال نماییم. درست است که میانگین نتایج ۱۰ مطالعه اولیه می‌تواند ارزشمند باشد ولی خوب یک نقطه ضعف نیز دارد. آیا نتایج همه مطالعات هم ارزش هستند؟ مسلماً این چنین نیست. بعضی مطالعات وسیعتر بوده و بعضی کوچکتر هستند. طبیعی است که بایست به نتایج مطالعات بزرگتر بیشتر توجه نمود. چاره این مشکل نیز ساده است؛ بهتر

است بجای میانگین ساده، میانگین وزن داده شده (weighted mean) را محاسبه نمود.

برای درک آسانتر میانگین وزن داده شده، شاید ملموسترین مثال برای افراد دانشجویی در ایران، نحوه محاسبه نمره کنکور ورود به دانشگاه باشد که در آن برای دروس مهمتر ضریب بیشتری قایل میشوند. در اینجا نیز بایست برای نتایج مطالعات بزرگتر (و یا دقیقتر) وزن بیشتری قائل شد.

حال اگر بخواهیم این مفهوم را به زبان ریاضی نشان دهیم میتوان فرمول آن را به شرح زیر درآورد:

$$\text{weighted mean} = \frac{\text{sum of (estimate * weight)}}{\text{sum of weights}} = \frac{\sum T_i W_i}{\sum W_i}$$

در این فرمول  $T_i$  برابر میزان اثر مشاهده شده در مطالعه  $i$  (در این مثال میزان اثر محافظتی مصرف سبزی به درصد در هر مطالعه) و  $W_i$  وزن آن مطالعه می باشد. در معادله فوق اگر همه مطالعات هم وزن باشند، مقدار میانگین وزن داده شده دقیقاً برابر میانگین ساده محاسبه خواهد بود، ولی اگر نتایج مطالعات هم ارزش نباشند، میانگین محاسبه شده به نتایج مطالعات با وزن بیشتر نزدیکتر خواهد بود.

البته این شروع کار است و در مراحل بعدی بایست دامنه اطمینان و یا میزان عدم تجانس نتایج (heterogeneity) را نیز بررسی نمود که در مراحل بعد به تفصیل بیان خواهند شد.

بر اساس مثال فوق میتوان انجام مراحل یک متاآنالیز را به شرح زیر خلاصه نمود:

۱. محاسبه یک شاخص مشترک در تمام مطالعات اولیه. این شاخص در مثال فوق درصد کاهش خطر بروز

بدخیمی معده در اثر مصرف سبزی بوده است. باید دقت نمود که شاخص محاسبه شده در تمام مطالعات یکسان باشد. بعنوان مثال ترکیب نمودن ضریب خطر خام (crude RR) با ضریب خطر تصحیح شده برای بعضی متغیرهای مخدوشکننده (adjusted RR) مسلماً دارای ایراد است و نتایج متاآنالیز را زیر سوال میبرد (توضیحات بیشتر در فصول قبل آورده شده است).

۲. ترکیب شاخص محاسبه شده بر اساس نتایج مطالعات اولیه و محاسبه میانگین وزنداده شده. در این قسمت بایست دقت شود که دقیقترین شیوه آماری مورد استفاده قرار گیرد که توضیحات بیشتر در ادامه آورده خواهد شد.

۳. بررسی میزان عدم تجانس نتایج مطالعات اولیه. این بررسی با استفاده از روشهای گرافیکی و آزمونهای آماری ممکن است.

۴. اصلاح روشهای بکارگرفته شده بر اساس میزان عدم تجانس نتایج مطالعات اولیه و بررسی علل احتمالی ایجادکننده آن از طریق روشهای آماری

۵. راییه نتایج بصورت روشن و استفاده صحیح از نمودارهای مربوطه

## ۶-۶. مدل فیکس در مقایسه با مدل تصادفی<sup>۵</sup>

در مدل فیکس، فرض بر آن است که پارامتر<sup>۶</sup> مورد بررسی در تمام مطالعات اولیه مقدار حقیقی ثابتی داشته و

---

<sup>۵</sup> fixed versus random models

تفاوت بین مقادیر محاسبه شده در مطالعات مختلف، تنها بدلیل تصادف حاصل از نمونه‌گیری‌های مکرر (sampling variation) بوده است. در حالی که در مدل تصادفی فرض بر این است که پارامتر اندازه‌گیری شده خود توزیع نرمالی داشته و هر مطالعه مقداری از آن را اندازه‌گرفته است. لذا در مدل تصادفی، تفاوت‌های دیده شده در بین نتایج مطالعات اولیه دو ریشه دارند؛ یکی به دلیل تصادف حاصل از نمونه‌گیری‌های مکرر و دیگری به دلیل تغییرات تصادفی مقدار واقعی پارامتر در مطالعات مختلف.

برای روشنتر شدن تفاوت مدل‌های فیکس و تصادفی شاید یک مثال لازم باشد. فرض نمایید هدف یک مطالعه تعیین فشار خون مردم شهر کرمان باشد. طبیعی است که مقدار این متغیر در کل جامعه (پارامتر) نامشخص است. برای برآورد آن یک نمونه تصادفی ۱۰۰ نفری گرفته می‌شود و میانگین فشار خون سیستمیک آنها محاسبه می‌گردد. در این مطالعه مقدار این شاخص ۱۲۴ میلی‌متر جیوه بدست می‌آید. در قدم بعد با استفاده از روش‌های ساده آماری، می‌توان برآورد فاصله‌ای از پارامتر مورد نظر را بدست آورد.

حال اگر چنین مطالعه‌ای تکرار شود پرواضح است که مقدار میانگین محاسبه شده در مطالعات مختلف دقیقاً یکسان نخواهد بود. اگر همه نمونه‌ها کاملاً تصادفی از کل جامعه گزینش شده باشند می‌توان بیان نمود که تفاوت‌های مشاهده شده، صرفاً بدلیل شانس و تصادف و به زبان آماری بدلیل تاثیر نمونه‌گیری‌های متعدد بوده -

---

<sup>۱</sup> پارامتر (Parameter) به مقدار یک متغیر در کل جامعه گفته می‌شود در حالی که شاخص (Index) به مقدار آن متغیر در نمونه مورد بررسی اطلاق می‌گردد بعنوان مثال میانگین وزن هنگام تولد نوزادان ایرانی می‌تواند یک پارامتر محسوب شود و مقدار محاسبه این میانگین در نمونه نوزادان مورد بررسی شاخص می‌باشد. در تحقیقات، هدف اصلی تعیین پارامترها است، ولی در عمل بدلیل عدم امکان اندازه‌گیری پارامترها، شاخص‌ها محاسبه و از روی آنها، پارامترها تخمین زده می‌شوند.

است. در این حالت می‌توان اطلاعات مطالعات مختلف را به روش فیکس با یکدیگر ترکیب نمود.

در مثال فوق اگر نمونه‌های گرفته‌شده دقیقاً از یک جمعیت ثابت گرفته‌نشده باشند و بعنوان مثال بعضی از مطالعات بر روی خانمها و بعضی بر روی آقایان؛ و گروهی بر روی جوانان و گروهی بر روی پیران کار کرده باشند، طبیعی است که به راحتی قابل قبول نیست که بگوییم تفاوت‌های دیده‌شده صرفاً بدلیل شانس و تصادف در نمونه‌گیری‌های مختلف است. لذا تفاوت‌های دیده‌شده بایست به شکل دیگری و با استفاده از مدل‌های تصادفی ترکیب شوند. البته لازم است این اصل را همیشه به خاطر داشت که آیا مطالعات از نظر علمی قابل ترکیب کردن هستند و یا پارامتر مورد نظر در جمعیت‌های مورد مطالعه، بصورت ذاتی با یکدیگر متفاوت هستند؟

اگرچه بصورت کلی پاسخ به سوال فوق ذهنی (subjective) است، ولی یک اصل کلی می‌تواند کمک کننده باشد و آن این است که آیا می‌توان تفاوت‌های بین پارامتر مذکور را در جمعیت‌های مختلف با استفاده از یک توزیع نرمال کلان توجیه نمود؟ به عبارت دیگر می‌توان فرض نمود که پارامترهای مذکور در جامعه‌های مختلف خود از یک توزیع نرمال پیروی می‌کنند؟ در صورتی که پاسخ به سوال فوق منفی باشد می‌توان تلویحاً برداشت نمود که ترکیب نمودن نتایج این مطالعات خالی از اشکال نیست، در غیر این صورت استفاده از مدل تصادفی قابل دفاع خواهد بود.

با توجه به این توضیحات و به زبان ریاضی می‌توان مدل‌های فیکس و تصادفی را به شکل زیر نوشت:

$$\hat{y} = \frac{\sum w_i \hat{y}_i}{\sum w_i} \quad \text{with} \quad w_i \approx \frac{1}{v_i}$$

$\hat{y}_i$  = مقدار شاخص محاسبه شده در مطالعه  $i$   
 $v_i$  = واریانس شاخص در مطالعه  $i$   
 $\hat{y}$  = مقدار برآورد نقطه‌ای شاخص بعد از ترکیب  
 نتایج مطالعات اولیه

فرمول فوق نشان میدهد که مقدار شاخص برآورد شده برابر مقدار وزن داده‌شده شاخص‌های محاسبه شده در مطالعات مختلف است. وزن اعمال شده هر مطالعه، در این فرمول نسبت معکوس با مقدار واریانس شاخص در آن مطالعه دارد. البته در صورت استفاده از روش *inverse variance* این تناسب به تساوی تبدیل شده و وزن هر مطالعه دقیقاً برابر معکوس واریانس خواهد بود.

در مدل فیکس فرض بر این است که  $\hat{y}_i$  از یک توزیع نرمال پیروی می‌کند به عبارت دیگر

$$\hat{y}_i \approx N(y, v_i)$$

یعنی مقدار مشاهده شده شاخص در هر مطالعه، از توزیع نرمالی با مقدار میانگین ثابت (پارامتر در کل جامعه،  $y$ ) و واریانسی برابر  $v_i$  پیروی می‌کند. در مدل تصادفی می‌توان بیان نمود که

$$\hat{y}_i \approx N(y_i, v_i) \quad \text{and} \quad y_i \approx N(y, \tau^2)$$

یعنی مقدار مشاهده شده شاخص در هر مطالعه از توزیع نرمالی پیروی می‌کند که میانگین آن برابر با پارامتر مذکور در همان جامعه ( $y_i$ ) است و پارمترهای مذکور در جامعه‌های مختلف خود از توزیع نرمالی پیروی می‌کنند که میانگین آن برابر با پارامتر مذکور در کل جوامع ( $y$ ) و واریانسی برابر  $\tau^2$  (با تلفظ *Tau*) می‌باشد. در نتیجه میزان تفاوت‌های دیده‌شده دارای دو منشأ ( $v_i$  و  $\tau^2$ ) است. در قسمتهای بعد ذکر خواهد شد که چگونه می‌توان مقدار  $\tau^2$  را برآورد نمود.



## ۶-۷. روشهای آماری رایج برای ترکیب نمودن شاخصها

متناسب با نوع شاخص مد نظر متا آنالیز، روش های متفاوتی برای ترکیب و تخمین شاخص تجمعی و دامنه اطمینان ۹۵٪ آن وجود دارد که در ادامه به مهمترین آنها اشاره و نقاط قوت و ضعف این روش ها شرح داده خواهد شد.

### روش معکوس واریانس (Inverse Variance Weighted Method)

همانگونه که در قسمت ۶-۵ و ۶-۶ شرح داده شده است یکی از ساده ترین شیوه ها برای ترکیب شاخص های میانگین وزن داده شده آنهاست و آسانترین شیوه وزن دادن به آنها، معکوس واریانس شاخص ها در مطالعات مختلف می باشد. به عبارت دیگر، در صورتی که در یک مطالعه واریانس محاسبه شده ۵ و در مطالعه دوم برابر ۱۰ باشد وزن مطالعه اول دو برابر وزن مطالعه دوم خواهد بود (یک پنجم در مقایسه با یک دهم). این شیوه برای محاسبه واریانس شاخص تجمعی نهایی و تعیین دامنه اطمینان آن نیز کاربرد دارد که فرمول محاسبه آن در قسمت ۶-۶ بیان شده است.

از محاسن این شیوه، درک آسان و محاسبه راحت آن می باشد. به شکلی که حتی با استفاده از برنامه های ساده ای مانند MS- excel نیز می توان محاسبات مذکور را در مدلهای فیکس تعیین نمود. همچنین این شیوه قابل استفاده برای ترکیب کلیه انواع شاخصها می باشد. به

عبارتی می‌توان در مطالعات توصیفی و تحلیلی از آن بهره جست. همچنین با کمی تعدیل می‌توان از این روش در مدل‌های تصادفی نیز استفاده نمود.

روش محاسبه واریانس و همچنین محاسبه شاخص‌های ساده‌ای مانند میانگین، فرمول مشخص و ثابت دارند ولی در خصوص شاخص‌هایی مانند ضریب خطر (Risk ratio) و نسبت خطر (Odds Ratio)، فرمول‌های متفاوتی وجود دارد که در ادامه به شیوه‌هایی اشاره می‌شود که می‌توان نسبت خطر (ضریب خطر) و واریانس آن را محاسبه نمود.

### روش Woolf

این روش برای محاسبه واریانس نسبت خطر (OR) استفاده می‌شود و کاربرد آن بسیار وسیع است. در این روش، واریانس نسبت خطر هر مطالعه، بر اساس جدول و فرمول زیر مربوطه محاسبه می‌شود.

نتیجه	عامل خطر	
	منفی	مثبت
منفی	a <sub>i</sub>	b <sub>i</sub>
مثبت	c <sub>i</sub>	d <sub>i</sub>

$$OR_i = \frac{a_i \times d_i}{b_i \times c_i}$$

در فرمول بالا نسبت خطر برای مطالعه  $i$  برابر است با حاصل ضرب تعداد افراد خانه‌های منفی و مثبت برای عامل خطر و نتیجاً به صورت همزمان تقسیم بر حاصل ضرب تعداد افرادی که تنها برای یکی از عوامل خطر و یا نتیجه مثبت بوده‌اند.

برای ترکیب نسبت خطرها در مطالعات مختلف می‌توان از فرمول زیر استفاده نمود.

$$OR = \frac{\sum (w_i \times OR_i)}{\sum w_i}$$

که در این فرمول  $w_i$  برابر وزن مطالعه شماره  $i$  می‌باشد و برابر معکوس واریانس هر مطالعه می‌باشد. برای محاسبه وزن هر مطالعه می‌توان از فرمول زیر استفاده نمود.

$$w_i = \frac{1}{Var(OR_i)} = \left( \frac{1}{\left( \frac{1}{a_i} + \frac{1}{b_i} + \frac{1}{c_i} + \frac{1}{d_i} \right)} \right)$$

البته ایراداتی نیز به این شیوه وارد است که شاید مهمترین آنها در زمانی باشد که حجم نمونه یک یا چند مطالعه وارد شده در متاآنالیز کم باشد و یا بعضی از خانه‌ها صفر باشند. مثلاً اگر برای تعیین نسبت خطر بین مصرف هروئین و سرطان معده، تعداد مبتلایان به سرطان در گروه مصرف‌کننده (خانه  $d$ ) در یک یا چند مطالعه صفر باشد، محاسبه نسبت خطر (OR) در آن مطالعات غیر ممکن است و لذا به سادگی نمی‌توان نتایج آنها را در

متاآنالیز وارد کرد. البته در اینگونه موارد، بعضی توصیه می کنند که يك عدد كوچك مانند ۰/۵ به تمام خانه هاي جدول اضافه شود تا امکان محاسبات فراهم گردد. حتي بعضي از نرم افزار هاي آماری مانند Rev.Man به صورت اتوماتيك به خانه هاي دارای صفر عدد ۰/۵ را اضافه مي کند. (در قسمتهای بعد این تکنیک به تفصیل شرح داده خواهد شد). با این وجود اگرچه اضافه نمودن یک عدد ثابت امکان محاسبه را فراهم می کند ولی در بعضی موارد دقت آن زیر سوال است. لذا در این گونه موارد توصیه می شود تا از روشهای جایگزین استفاده گردد.

### روش Manteh – Haenzel

این شیوه تنها برای متاآنالیز متغیرهای وابسته دو حالتی (Dichotomous) کاربرد دارد. به عنوان مثال این شیوه را می توان در ترکیب نتایج مطالعاتی که بر روی مرگ و زندگی و یا بهبود و عدم بهبود بکار برد.

از محاسن این شیوه آن است که می تواند حتی در مواردی که حجم نمونه مطالعات کم هستند و یا بعضی خانه های جداول صفر می باشند نیز با دقت نسبتاً بالا استفاده نمود.

مقدار نسبت خطر حاصل از متاآنالیز بر اساس این روش، مشابه روش قبلی است و فرمول کلی آن برابر خواهد بود با

$$OR = \frac{\sum (w_i \times OR_i)}{\sum w_i}$$

ولی تفاوت اصلی آن با روش قبلی در تعیین وزن مطالعات است و فرمول کلی عبارت است از

$$w_i = \frac{b_i c_i}{N_i}$$

که  $N_i$  یعنی حجم نمونه هر مطالعه و روش محاسبه آن برابر است با

$$N_i = a_i + b_i + c_i + d_i$$

از ترکیب و خلاصه نمودن فرمولهای فوق میتوان به فرمول زیر دست یافت:

$$OR = \frac{\sum (a_i d_i / N_i)}{\sum (b_i c_i / N_i)}$$

اگرچه استفاده از روش مذکور در اکثر موارد برای ترکیب OR بکار میرود، ولی میتوان این روش را برای ترکیب نمودن ضریب خطر (RR) نیز بکار برد.

از نقاط ضعف این روش آن است که متأسفانه فقط برای مدل های فیکس میتوان از آن بهره جست و استفاده از آن برای مدل های تصادفی ممکن نیست.

**Peto** روش

این شیوه نیز مانند قبل فقط برای مواردی کاربرد دارد که متغیر وابسته دو حالتی است و هدف، ترکیب نسبت خطر (OR) می‌باشد. در این شیوه مقدار OR بر اساس تفاوت مقادیر مشاهده شده (Observed) و مورد انتظار (expected) تخمین زده می‌شود.

$$\log(OR_i) = \frac{a_i - \exp(a_i)}{\text{Var}(a_i)}$$

در این فرمول مقدار مورد انتظار برای  $a_i$  برابر

$$\exp(a_i) = \frac{(a_i + b_i)(a_i + c_i)}{N_i}$$

و مقدار واریانس  $a_i$  برابر

$$\text{Var}(a_i) = \frac{(a_i + b_i)(a_i + c_i)(d_i + b_i)(d_i + c_i)}{N_i}$$

با توجه به فرمول‌های فوق مقدار نسبت خطر برآورد شده کل مطالعات برابر خواهد بود با :

$$\log(OR) = \frac{\sum (w_i \times \log(OR_i))}{\sum w_i} \text{ and } \text{Var}(\log(OR)) = \frac{1}{\sum w_i}$$

البته روش peto برای حالت‌هایی که حجم نمونه کم است و یا بعضی خانه‌های جدول صفر هستند نیز کاربرد دارد

ولی در اکثر اوقات تخمین نسبت خطر (OR) با فرمول پیشنهادی دارای سوگیری است و نتایج آن قابل اعتماد نیست بخصوص اگر مقدار OR محاسبه شده از یک، فاصله داشته باشد و به عبارتی نشان دهنده ارتباط قوی بین متغیر مستقل و وابسته نشان دهند. به همین دلیل کاربرد این شیوه در متاآنالیز بسیار محدود شده است و لذا از توضیحات بیشتر در مورد این روش خودداری می شود. لازم به ذکر است که این روش نیز فقط در مدل های فیکس کاربرد دارد.

البته باید ذکر شود که این روش برای تحلیل و ترکیب نتایج مطالعات بقاء (Survival analysis) نیز قابل استفاده است که شرح آن از حوزه این کتاب فراتر است.

## ۶-۸. روشهای آماری رایج برای ترکیب نمودن شاخصها در مدل های تصادفی

در مواردی که تفاوت های بین یافته های مطالعات مختلف بیشتر از حدی باشد که توسط احتمال شانس و تصادف حاصل از نمونه گیری های مکرر قابل توجیه گردد، ترکیب نتایج مطالعات با شیوه های یاد شده فوق دارای ایراد است و به عبارتی باید از مدل های تصادفی استفاده نمود که اجازه می دهند تا مقدار پارامتر مورد نظر در مطالعات مختلف تغییر کند (لطفاً به توضیحات مدل های تصادفی در قسمت ۶-۶ مراجعه شود).

در روش های یاد شده فوق، مقدار وزن مطالعات بر اساس معکوس واریانس آنها بدست می آید در حالی که در این روش علاوه بر عکس واریانس، میزان تفاوت شاخص محاسبه

شده در مطالعه از مقدار برآورد کلی شاخص نیز به عنوان جزئی از وزن مطالعه مد نظر خواهد بود. به همین دلیل مقدار محاسبه شده دامنه اطمینان با این روشها از مقادیر محاسبه شده در روشهای قبلی که در واقع روش های فیکس بودند وسیعتر خواهد بود.

با توجه به این توضیحات می توان به صورت کلی نوشت که برآورد نهایی شاخص  $Y$  برابر بود با

$$Y = \frac{\sum w_i^* y_i}{\sum w_i}$$

و مقدار  $w_i^*$  یعنی وزن هر مطالعه

$$w_i^* = \frac{1}{(v_i + \tau^2)}$$

در این روشها، سعی می شود مقدار  $\tau^2$  یعنی واریانس شاخص مذکور در بین مطالعات تخمین زده شود. افزایش مقدار  $\tau^2$ ، نشان دهنده افزایش تفاوتها بین مطالعات و به عبارتی افزایش عدم تجانس مطالعات (Heterogeneity) خواهد بود.

در واقع تفاوت بین روشهای مختلف مدل های تصادفی، در روش تخمین مقدار  $\tau^2$  می باشد که در ادامه به بیان مهمترین این شیوه ها به صورت مختصر و بدون پرداختن به روشهای آماری پیچیده، خواهیم پرداخت.



## روش Der Simonian and Laird

در این روش مقدار  $\tau^2$  بوسیله حل مطالعات نسبتاً پیچیده تخمین زده می‌شود و روش نسبتاً رایجی است. به جهت ساده نمودن موضوع از پرداختن به جزئیات خوداری می‌شود و در فصل آخر روش محاسبه با استفاده از نرم-افزار stata آموزش داده خواهد شد.

## روش متا-رگرسیون

کاربرد این روش نه تنها در تخمین  $\tau^2$  بکار می‌رود بلکه کمک می‌کند تا دلایل ایجاد عدم تجانس بین یافته‌های تحقیقات نیز مورد بررسی قرار گیرند. مبانی ذهنی این شیوه، کاملاً شبیه رگرسیون معمولی است. به عبارتی بایست متغیر وابسته را که در واقع همان شاخص آماری است که در تکتک مطالعات محاسبه می‌شود (مانند RR, OR و یا میانگین) برگزید. در صورتی که مدل متا-رگرسیون بدون هر گونه متغیر دیگری تنظیم شود مقدار  $\tau^2$  تخمین زده شده دقیقاً همان  $\tau^2$  است که در روش Der Simonian & Laird تعیین می‌شود. ولی حسن روش متا-رگرسیون آن است که می‌توان با وارد نمودن بعضی متغیرهای مستقل، میزان تغییرات  $\tau^2$  را سنجید. به عبارتی در صورتی که وارد نمودن بعضی مشخصات مطالعات بتواند  $\tau^2$  کاهش دهد می‌توان آن شاخصها را عامل عدم تجانس یافته‌های مطالعات دانست. به عنوان مثال اگر در خصوص تأثیر استرپتوکیناز بر کاهش خطر مرگ در اثر حمله حاد قلبی قصد متاآنالیز داشته باشیم و مقدار  $\tau^2$  محاسبه شده برابر ۲/۵ باشد، نشان می‌دهد که یافته‌های مطالعات با یکدیگر همخوانی کافی نداشته‌اند. مثلاً در بعضی از تحقیقات این درمان ۴۰٪ مؤثر و بعضی مطالعات حتی ۱۰٪

مضر دانسته شده است. حال می‌توان در مدل متارگرسیونی متغیرهایی را وارد نمود و بررسی کرد که آیا این متغیرها می‌توانند  $\tau^2$  را کاهش دهند یا خیر؟ فرض کنید با وارد نمودن میانگین سنی نمونه‌ها در مطالعات مختلف، مقدار  $\tau^2$  به  $1/2$  کاهش یابد، نتیجه‌گیری شما چه خواهد بود؟

بله درست فهمیدید، نشان می‌دهد که اثر بخشی استرپتوکیناز وابسته به سن است چرا که با حذف اثر سن، میزان عدم تجانس مطالعات ( $\tau^2$ ) کاهش قابل ملاحظه‌ای یافته است. حتی با نگاه کردن به نتایج جدول متارگرسیونی و دیدن علامت ضریب رگرسیونی مربوط به میانگین سنی می‌توان قضاوت نمود که جهت تأثیر سن چگونه بوده است؟ در صورتی که این ضریب مثبت باشد نشان می‌دهد که با افزایش میانگین سنی، OR افزایش می‌یابد و به عبارتی اثر درمانی استرپتوکیناز زیاد خواهد شد و در مقابل، علامت منفی این ضریب مبین اثر بخشی کمتر (کوچکتر شدن OR) استرپتوکیناز در سنین بالاتر است.

با این تکنیک حتی می‌توان در خصوص تفاوت‌های بین گروهی نیز قضاوت کرد. مثلاً اگر گذاشتن جنس (اگر تمام نمونه‌های مطالعات وارد شده فقط از یک جنس باشند) و یا درصد مرد بودن نمونه در هر مطالعه، مقدار  $\tau^2$  را تغییر قابل توجه‌ای ندهد می‌توان نتیجه گرفت که جنس، یک فاکتور مؤثر در ایجاد Heterogeneity نیست که تلویحاً نشان می‌دهد که اثر درمانی استرپتوکیناز در دو جنس احتمالاً متفاوت نیست.

برای محاسبه  $\tau^2$  در مدل‌های متارگرسیونی چهار تکنیک متفاوت وجود دارد که عبارت هستند از:

روش گشتاوری (Moment base) که سعی می کند از طریق حل معادله به جواب برسد و به همین دلیل در بعضی موارد نمی تواند به جواب برسد. از نظر عددی  $\tau^2$  محاسبه شده به این شیوه، دقیقاً برابر  $\tau^2$  روش Der Simonian Lairn است.

روش Maximum Likelihood که منطبق بر نام این روش، به روش Likelihood مقدار  $\tau^2$  برآورد می شود این روش تقریباً در اکثر موارد به جواب می رسد ولی گاه بدلیل شکل توزیع متغیرها عدد محاسبه شده برای  $\tau^2$  بسیار کوچک و یا بسیار بزرگ است.

روش Restricted Likelihood: شبیه به روش فوق است ولی برای اصلاح  $\tau^2$  محاسبه شده، معادله Likelihood بر روی توزیع پردازش شده (trimmed) رسم می شود.

روش بیزین (Bayesian): در این روش منطبق بر تئوری بیزین، مقدار اولیه (Prior) برای  $\tau^2$  و همچنین شاخص مورد نظر تعیین می شود و سپس بر اساس نتایج و منحنی Likelihood، آن مقدار تصحیح شده (Posterior) محاسبه می گردد.

اگر چه تفاوت های بین نتایج این روش به خصوص در تخمین  $\tau^2$  قابل ملاحظه است ولی متأسفانه مطالب نوشته شده در خصوص مزایا و یا معایب این شیوه ها بسیار محدود است. به اعتقاد نگارنده، شاید در اکثر موارد روش اول (گشتاوری) و روش سوم (Restricted likelihood) روشهای ارجح باشند.

۶-۹- خلاصه روش های آماری متناسب با نوع شاخص

احتمالاً در قسمت های قبلی با اسامی متعدد و روش های آماری مختلف آشنا شده اید ولی به جهت مرور و همچنین ارایه يك خلاصه از توضیحات قبلی سعی شده است، اهم کاربرد روش های مختلف متاآنالیز در جدول زیر خلاصه شود.

نوع داده	شاخص آماری	روش آماری نوع
دو حالته	نسبت خطر (Odds Ratio)	Woolf فیکس
		Mantel-Haenzel فیکس
		Peto فیکس
دو حالته	ضریب خطر (Risk Ratio)	Der simonian-Laird تصادفی
		Meta-Regression تصادفی
		Mantel-Haenzel فیکس
دو حالته	تفاضل خطر (Risk Difference)	Inverse Variance فیکس
		Der Simonian-laird تصادفی
		Meta-Regression تصادفی
عددی	تفاضل میانگین ها و یا تفاضل میانگین های استاندارد	Inverse Variance فیکس Der Simonian-laird تصادفی Meta-Regression تصادفی

## ۶-۹. مفهوم عدم تجانس (Heterogeneity)

هرگونه عدم یکنواختی در مطالعات وارد شده در متاآنالیز ( $\tau^2$ ) اصطلاحاً عدم تجانس خوانده می شود. عدم یکنواختی ممکن است منشا بالینی داشته باشد که مهمترین دلایل آن، تفاوت در نمونه های مورد مطالعه، تفاوت در نوع درمانها و تفاوت در متغیرهای وابسته است. مثلاً ممکن است که نمونه های وارد شده در مطالعات از نظر سنی و جنسی و یا شدت بیماری مورد مطالعه متفاوت بوده باشند و یا دوز دارو یا نحوه تجویز در مطالعات متفاوت بوده باشد و یا متغیر وابسته با روش ها و دستگاه های مختلف سنجیده شده باشد.

دلیل دوم عدم تجانس ممکن است منشا متدلوژیک داشته باشد و بدلیل شیوه های مختلف مطالعه ایجاد شده باشند. طبیعتاً مطالعات مشاهده ای و یا مداخله ای از نظر مقدار، شاخصهای نهایی ممکن است متفاوت باشند. همچنین طول مدت متفاوت پیگیری و یا شیوه حذف اثر فاکتورهای مخدوش کننده می تواند منشا عدم تجانس باشد.

به هر شکل هرگونه عدم تجانس چه به شکل متدلوژیک و چه به شکل بالینی، باعث بروز عدم تجانس آماری (Statistical heterogeneity) می شود.

به هر شکل در صورتی که میزان عدم تجانس زیاد باشد و به عبارتی مشخصات بالینی و متدلوژیک مطالعات بسیار متفاوت باشند، انجام متاآنالیز و ترکیب نتایج آنها خالی از اشکال نیست.

### نحوه اندازه گیری میزان عدم تجانس

برای محاسبه میزان عدم تجانس آزمون ساده کای-دو انجام می شود. مقدار این آماده با استفاده از فرمول زیر محاسبه می شود. لازم به ذکر است که درمتا-آنالیز کای - دو با علامت Q نشان داده می شود.

$$Q = \sum w_i (y_i - \bar{y})^2$$

$$df=k-1$$

به عبارتی دیگر برای محاسبه Q باید تفاضل مقدار شاخص محاسبه شده در هر مطالعه از برآورد کلی آن (حاصل متاآنالیز) به توان دو رسیده و در وزن مطالعه مذکور ضرب شود. حاصل جمع این مقادیر برابر Q خواهد بود و درجه آزادی این آماره برابر تعداد مطالعات منهای یک می باشد.

بر این اساس می توان با کمک از جدول کای - دو، مقدار P-value را محاسبه نمود. کوچک بودن مقدار P-value مبین تفاوت معنی دار نتایج مطالعات و به عبارتی هتروژنیته است.

البته با توجه به اینکه معمولاً تعداد مطالعات وارد شده در یک متاآنالیز کم و محدود است لذا توان (Power) این آزمون نیز زیاد نمی باشد و به همین دلیل در اکثر موارد سطح معنی داری، ۰/۱ در نظر گرفته می شود نه ۰/۰۵

شاخص دیگری که در اینجا مورد استفاده قرار می‌گیرد  $I^2$  است که به این شکل محاسبه می‌شود.

$$I^2 = \left( \frac{Q - df}{Q} \right) \times 100$$

این شاخص در سال ۲۰۰۳ توسط Higgins معرفی شده و بیان می‌کند که چند درصد تفاوت‌های مشاهده شده بین شاخص‌های مطالعات متفاوت به دلیل عدم تجانس مطالعات است تا به دلیل اثر نمونه‌گیری‌های مکرر (Sampling error) به صورت کلی عدد بالای ۵۰٪ مبین وجود عدم تجانس قابل ملاحظه در بین مطالعات است.

البته با نگاه کردن به نمودار مربوط به دامنه اطمینان شاخص مورد نظر در مطالعات مختلف نیز می‌توان در خصوص هتروژنیستی نظر داد. عدم همپوشانی قابل ملاحظه دامنه‌های اطمینان ۹۵٪ شاخص مطالعات مختلف، خود دلیلی بر وجود عدم تجانس بین مطالعات است.

### شیوه برخورد با عدم تجانس آماری مطالعات

در صورت مشاهده عدم تجانس قابل ملاحظه، در بین نتایج مطالعات مختلف باید یکی از استراتژی‌های زیر را اتخاذ نمود.

۱. بررسی مجدد مطالعات برای کسب اطمینان از صحت نتایج و اصلاح هر گونه ایراد و یا خطا در هنگام استخراج، ثبت و یا ورود اطلاعات به رایانه

۲. صرف نظر کردن از متاآنالیز: همانگونه که قبلاً نیز بیان شد در صورت وجود عدم تجانس بسیار زیاد مطالعات انجام متاآنالیز غیر قابل قبول است و بایست حداکثر به توصیف نتایج و اخذ روش های کیفی یعنی meta-Synthesis اکتفا کرد.

۳. بررسی ریشه ایجاد کننده عدم تجانس از طریق انجام تحلیل در زیر گروه ها: برای این امر بایست مطالعات را بر اساس بعضی از فاکتورهای احتمالی ایجاد کننده عدم تجانس تقسیم نمود و در هر دسته تحلیل آماری مجزا انجام داد. مثلاً می توان مطالعات را بر اساس نوع آنها (مشاهده ای و یا مداخله ای) و یا بر اساس گروه سنی نمونه ها تقسیم نمود و سپس در هر زیر گروه متاآنالیز مجزا انجام داد. در صورتی که در هر زیر گروه عدم تجانس کاهش قابل ملاحظه ای یافته شود، مبین آن است که عامل ایجاد کننده عدم تجانس، فاکتور مذکور بوده است. این روش در عین سادگی یک ایراد اساسی دارد و آن مشکل حجم نمونه است. چرا که در اکثر متا-آنالیزها تعداد مطالعات قابل دسترسی کم می باشد و دسته بندی همین تعداد محدود مشکل ساز خواهد بود.

۴. عدم توجه به وجود عدم تجانس و انجام متاآنالیز به روش فیکس که البته در این منابع اصلاً این شیوه توصیه نمی شود.

۵. انجام متاآنالیز به روش تصادفی: اگرچه در تعدیل برآورد مربوطه و دامنه اطمینان کمک می کند ولی بدلیل عدم کارایی این روش در تعیین دلیل احتمالی ایجاد کننده عدم تجانس در بسیاری از مواقع بهترین شیوه نیست.



۶. تغییر مبنای شاخص: گاهی اوقات با تغییر شاخص مورد نظر می‌توان عدم تجانس را کاهش داد. مثلاً می‌توان بررسی کرد که آیا میزان عدم تجانس برای ضریب خطر و نسبت خطر و یا تفاضل میانگین‌ها و تفاضل استاندارد میانگین برابر است؟

۷. حذف بعضی از مطالعاتی که بیشترین عدم تجانس را با بقیه دارند: البته از این روش نباید به راحتی استفاده نمود. ابتدا بایست با دقت دلیل احتمالی تفاوت یک یا چند مطالعه با بقیه را سنجید. به واقع حذف بعضی مطالعات بدون دلیل و یا شناخت علت تفاوتها باعث ایجاد سوگیری در نتایج شده و یافته‌های متاآنالیز را زیر سؤال خواهد برد. به عنوان مثال فرض کنید؛ هفت مطالعه نشان دهند که استوپتوکیناز میزان بقای بعد از MI را افزایش می‌دهد ولی دو مطالعه عکس آن را نشان دهند. طبیعی است این دو مطالعه باعث افزایش عدم تجانس به میزان زیاد خواهد شد. ولی بدون دانستن دلیل عدم همسانی نتایج این دو مقاله با بقیه اصلاً صحیح نیست تا این مطالعات را حذف و بر روی ۷ مطالعه باقیمانده متاآنالیز انجام داد.

۸. انجام متارگرسیون: در این شیوه همانگونه که قبلاً نیز شرح داده شد؛ نه تنها امکان بررسی وجود عدم تجانس بین یافته‌های مطالعات است بلکه کمک می‌کند تا بدانیم دلیل ایجاد عدم تجانس چیست؟ یعنی همانند روش سوم (تحلیل زیر گروه‌ها) نه تنها ریشه عدم تجانس را بررسی می‌کند بلکه مشکل کم بودن حجم نمونه را تا حدود زیادی برطرف نموده و میزان و جهت اثر فاکتورهای تأثیرگذار را نیز روشن می‌کند. به

همین دلیل این روش یکی از بهترین شیوه‌های  
برخورد با عدم تجانس است.

---

## ۷. نمودار های اصلی در متاآنالیز و آزمونهای آماری مرتبط

---

نمودارها در متاآنالیز همانند کاربرد آنها در سایر تحلیلها به بیان ساده تر نتایج کمک می کند و به سادگی مفاهیم پیچیده را به نمایش می گذارند. کاربرد این نمودارها به حدی است که تقریباً در تمامی مقالات متاآنالیزی حداقل یکی از نمودارهای اصلی نمایش داده شده اند. در این فصل سعی خواهد شد سه نمودار اصلی و رایج در متاآنالیز شرح داده شوند و در فصل نهم با نحوه رسم آنها در برنامه Stata آشنا خواهید شد.

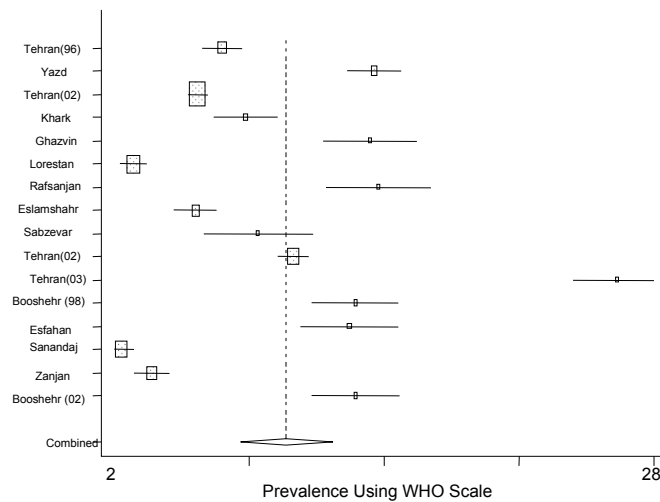
### ۷-۱. نمودار انباشت (Forest Plot)

این نمودار رایج ترین نوع نمودار در متاآنالیز است که اطلاعات تک تک مطالعات و برآیند نهایی آنها را به نمایش می گذارد. شکل کلی این نمودار در ذیل نمایش داده شده است. در این نمودار فراوانی دیابت به تفکیک استان های کشور به نمایش گذاشته شده است. هر مربع مقدار فراوانی دیابت را در هر استان نمایش می دهد و اندازه آن متناسب با وزنی است که آن مطالعه در متاآنالیز داشته است. خط افقی اطراف هر مربع نیز دامنه اطمینان ۹۵٪ را نشان می دهد.

گاهی اوقات دو خط عمودی نیز رسم می شود. یکی از این خطوط منطبق بر محور عمودی لوزی است و نشان می دهد

که در کل مقدار اثر بیان شده در متاآنالیز چقدر است؟

خط عمودی دوم معمولاً بر روی عددی رسم می شود که نشان دهنده مقدار فرضیه صفر است، مثلاً برای مواردی که نسبت خطر و یا ضریب خطر نمایش داده شده است عدد مذکور یک می باشد یعنی عددی که مبین عدم ارتباط متغیر مستقل و وابسته است. عدد مذکور برای تفاضل خطرات برابر صفر است و نشان می دهد که بر اساس فرضیه صفر خطر در گروه های مختلف مورد مطالعه برابر فرض شده است.



نمودار ۲: نمودار انباشت میزان شیوع دیابت بر اساس معیار سازمان بهداشت جهانی در استانهای مختلف ایران؛ لوزی پایین نمودار میزان پیش‌بینی شده بر اساس مدل تصادفی را نشان می‌دهد.

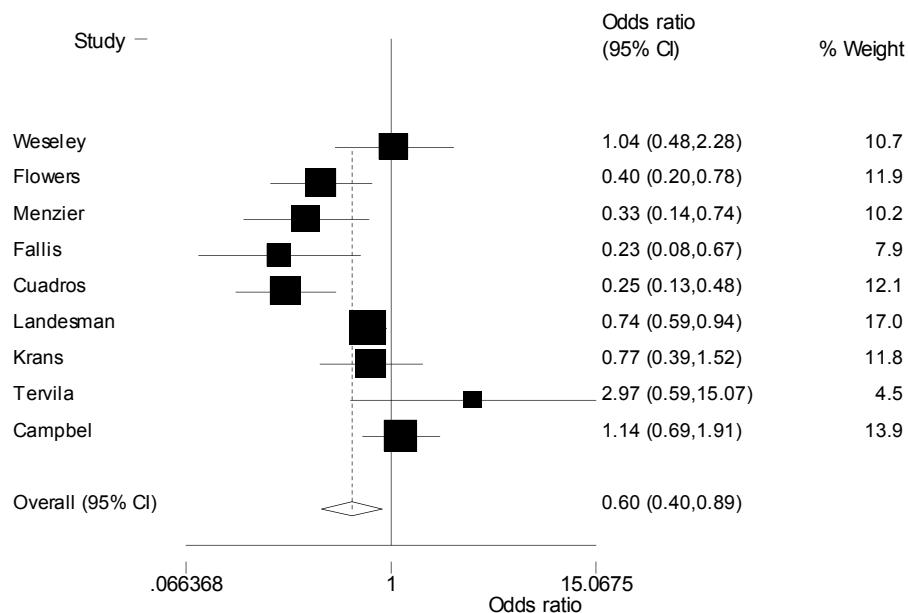
در پایین نمودار انباشت یک لوزی رسم می شود که در واقع مقدار برآیند وزن داده شده مربع ها است. قطر افقی لوزی نیز نشان دهنده دامنه اطمینان ۹۵٪ برآیند

شاخص مورد مطالعه است. در این شکل مقدار متناظر با مرکز لوزی برابر ۴.۶٪ است که نشان می‌دهد تخمین کلی دیابت در کشور برابر با عدد مذکور است و دامنه اطمینان ۹۵٪ آن نیز از ۴٪ تا ۵.۶٪ می‌باشد.

لازم به ذکر است که محور افقی در زمانی که شاخص مورد مطالعه یک نسبت است (مانند Hazard Ratio, Rate Ratio, Risk Ratio, Odds Ratio) به صورت لگاریتمی نمایش داده می‌شود.

در زیر نمودار انباشت مربوط به متا-آنالیزی را نشان می‌دهد که بر روی اطلاعات مربوط به تاثیر مصرف دیورتیک بر پیشگیری از بروز پره‌اکلامپسی انجام شده است. در این نمودار محور افقی نسبت خطر (OR) را نمایش می‌دهد و مقیاس آن لگاریتمی است. در کنار اسم هر مطالعه اسم نویسنده اول؛ نسبت خطر مشاهده شده و وزن مطالعه بر است روش تصادفی نمایش داده شده است. همان گونه که ملاحظه می‌شود بر روی عدد نسبت خطر یک، خط عمودی تیره‌ای کشیده شده است تا به خواننده کمک کند که نتایج کدام مطالعه معنی‌دار است. اگر خط افقی مربوط به دامنه اطمینان مطالعه‌ای این خط را قطع کند بدین معنی است که از نظر آماری معنی‌دار نیست. همچنین بر روی محور عمودی لوزی پائین نمودار نیز خط عمودی کشیده شده است که نشان می‌دهد نسبت خطر پیش‌بینی شده بر اساس متا-آنالیز به روش تصادفی برابر با ۰.۶ است و دامنه اطمینان آن نیز عدد یک را در بر نمی‌گیرد پس می‌توان نتیجه گرفت که مصرف دیورتیکها به صورت معنی‌داری بروز پره‌اکلامپسی را کاهش می‌دهد.

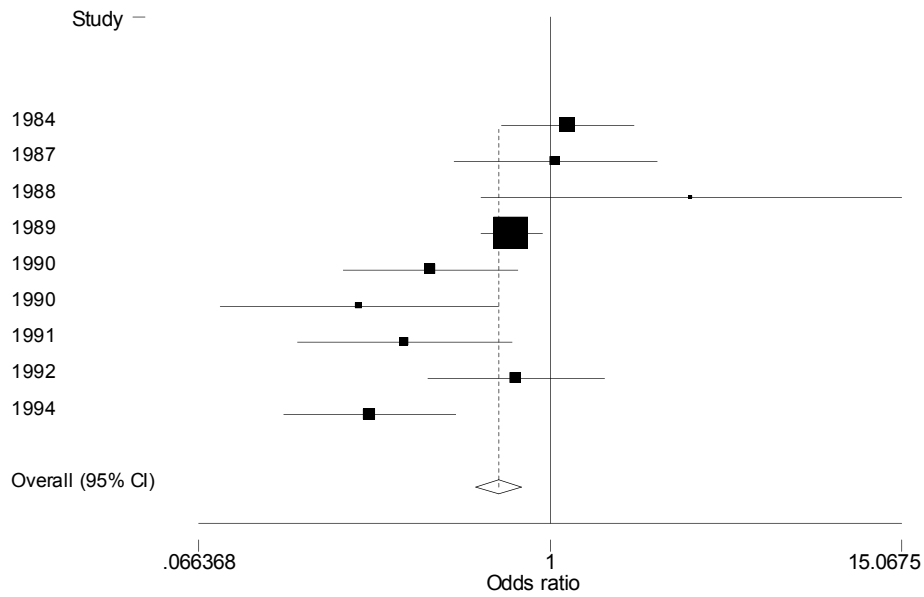
## جستجوی ساختاریافته و متاآنالیز



نمودار 3: نمودار انباشت نسبت خطر مصرف دیورتیک بر بروز پره-اکلامپسی

ترتیب اطلاعات مطالعات در محور عمودی نمودار انباشت قابل تنظیم است و بهتر است یک روند منطقی داشته باشد. گاهی اوقات مطالعات را بر اساس سال انتشار مرتب می‌کنند و گاهی اوقات بر اساس حجم نمونه. به هر شکل بهترین معیار مرتب کردن، معیاری است که احتمالاً بیشترین تأثیر را در ایجاد هتروژنیته داشته است. در صورت مؤثر بودن متغیر مرتب کننده در ایجاد عدم تجانس انتظار می‌رود که از بالا به پایین مربع‌ها نمایش داده شده در نمودار تقریباً در یک روند مشخص از راست به چپ و یا برعکس تغییر مکان دهند. به عنوان نمونه در متا-آنالیز نسبت خطر مصرف دیورتیک در پیشگیری از ابتلا به پره‌اکلامپسی، در صورتی که مطالعات بر اساس سال تحقیق مرتب شوند، به وضوح یک روند در نتایج تحقیقات بر اساس سال تحقیق دیده می‌شود

و مطالعات جدید تاثیر بیشتری بر اثربخشی دیورتیکها نشان می‌دهد.



نمودار:4 نمودار انباشت نسبت خطر مصرف دیورتیک بر بروز پره-اکلامپسی، تفاوت این نمودار با نمودار قبل در آن است که مطالعات بر اساس سال تحقیق مرتبط شده‌اند و نشان می‌دهد که مطالعات سالهای اخیر اثربخشی بیشتری از مصرف دیورتیکها نشان می‌دهد.

بر اساس نتایج بدست آمده از نمودار انباشت فوق می-توان دریافت که احتمالاً سال تحقیق می‌تواند درصد قابل توجهی از هتروژنیته بین نتایج مطالعات را توجیه می‌کند و بهتر است این متغیر را به عنوان یک متغیر تاثیرگذار در متا-رگرسیون وارد نمود.

## ۲-۷. نمودار کیفی (Funnel Plot) برای بررسی شانس بروز سوگیری در انتشار نتایج

پر واضح است که دقت مطالعات بزرگتر در برآورد پارامتر مورد نظر بیشتر از مطالعات کوچکتر است. به عبارتی خطاهای تصادفی در مطالعات بزرگتر کمتر وجود دارند و لذا احتمالاً برآورد پارامتر مورد نظر در مطالعات بزرگ به یکدیگر نزدیکتر است. به زبان دیگر انتظار می‌رود که شاخص مورد اندازه‌گیری در مطالعات بزرگ نزدیک یکدیگر باشد.

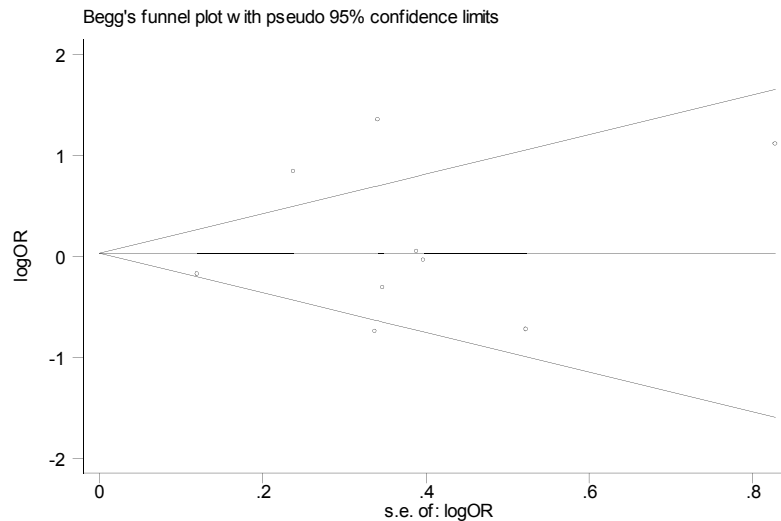
به عنوان مثال فرض کنید، جهت برآورد رابطه بین مصرف دیورتیک و خطر پره‌اکلامپسی چندین تحقیق صوت گرفته‌باشد. مشخص است که هدف اصلی تمام این مطالعات یافتن پارامتر مورد نظر مثلاً نسبت خطر (Odds Ratio) واقعی بین مصرف دیورتیک و پره‌اکلامپسی است. واقعیت آن است که تقریباً همیشه مقدار این پارامتر برای ما یک مجهول است و برای تخمین زدن آن تحقیق می‌کنیم. اما فرض کنید که ما مقدار این پارامتر دقیقاً برابر  $0/6$  باشد، به عبارت دیگر واقعاً مصرف دیورتیک، خطر اکلامپسی را حدوداً  $40\%$  کم کند. بنابراین انتظار داریم تک تک مطالعات انجام شده این عدد  $0/6$  را برآورد نمایند. مسلماً در اکثر تحقیقات نسبت خطر بدست آمده دقیقاً  $0/6$  نیست ولی در صورتی که مطالعات از نظر متدلوژی تقریباً صحیح انجام شده باشند؛ انتظار داریم نسبت خطر محاسبه شده در مطالعات بزرگتر به عدد  $0/6$  نزدیکتر بوده و به عبارتی پراکندگی تصادفی نسبت خطر مطالعات بزرگتر کمتر باشد. ولی در مقابل تنها به دلیل شانس و تصادف تخمین بدست آمده در مطالعات کوچکتر از  $0/6$  فاصله بیشتری خواهند داشت.



حال برای نمایش این واقعیت سعی کنید نموداری را رسم نمائید که در یکی از محورهای خود (معمولاً محور Y) مقدار شاخص مورد نظر مانند (چون مدل از نوع مدلهای **multiplicative** است پس بجای خود نسبت خطر باید لگاریتم آن استفاده شود) و در محور دیگر انحراف استاندارد شاخص مذکور (باز دقت شود که در این مثال باید از لگاریتم انحراف معیار استفاده شود) را نمایش می دهد. سپس با توجه به شاخص مورد نظر و انحراف استاندارد آن در هر مطالعه، نقطه متناظر در نمودار نمایش داده می شود. در این نمودار مطالعات دقیقتر دارای انحراف استاندارد کمتری بوده و لذا بر روی محور X که Se را نشان می دهد عدد کوچکتری را به خود اختصاص می دهند و در سمت چپ نمودار و نزدیک صفر محور عمودی قرار خواهند داشت. در مقابل مقادیر مربوط به مطالعات کوچکتر که خطای معیار بیشتری دارند در سمت راست نمودار واقع خواهند شد.

بر اساس توضیحات فوق، انتظار بر آن است که تخمینهای رایج شده از مطالعات بزرگتر و با دقت بیشتر به یکدیگر نزدیکتر و در مقابل تخمینهای مطالعات کوچکتر از یکدیگر بیشتر فاصله بگیرند.

در نتیجه بنظر می رسد که پراکندگی نقاط باید در این نمودار شکل قیف خوابیده را به خود بگیرد و دهانه تنگ قیف در سمت چپ یعنی محل استقرار نقاط مربوط به مطالعات دقیقتر و دهانه گشادتر قیف در سمت راست یعنی محل استقرار نقاط مربوط به مطالعات کوچکتر باشد.



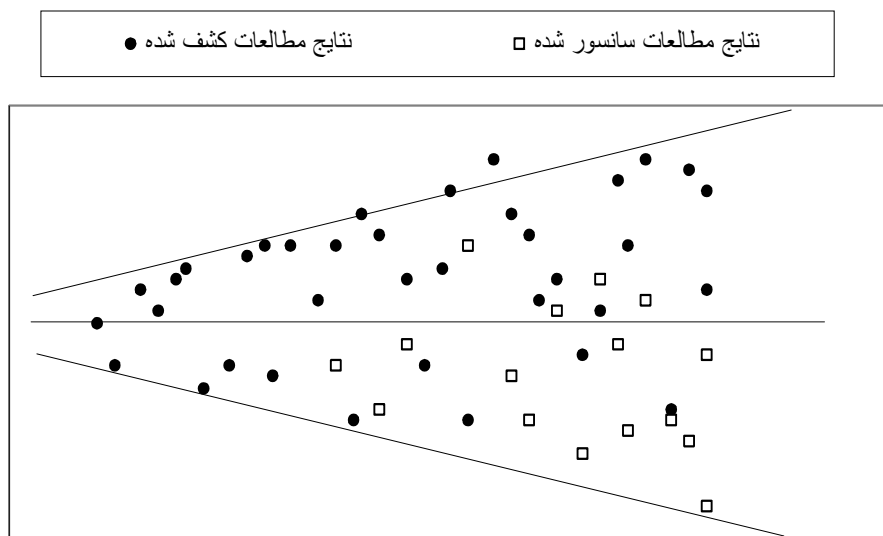
نمودار ۵: نمودار کیفی نتایج مربوط به نسبت خطر بین مصرف دیورتیک و ابتلا به پره اکلامپسی در دوران بارداری؛ این نمودار معمولاً برای نشان دادن میزان احتمال بروز سوگیری در انتشار نتایج بکار می‌رود.

در صورت وجود سوگیری در چاپ، انتظار بر آن است که امکان دستیابی به نتایج بعضی تحقیقات وجود نداشته باشد و در نتیجه امکان وارد نمودن اطلاعات این مطالعات در نمودار کیفی ممکن نیست. به عبارتی بعضی نقاط نمودار کیفی، به عنوان نقاط سانسور شده از نظر دور خواهند ماند.

با این پیش فرضها، می‌توان پیش‌بینی نمود که نتایج بعضی مطالعات شانس کمتری برای چاپ دارند، و البته خیلی دور از ذهن نیست که به پذیریم شانس چاپ مطالعات کوچکتر به مراتب کمتر از مطالعات بزرگتر است که احتمالاً از نظر نوع طراحی نیز دقیقتر بوده‌اند.

پس در صورت وجود سوگیری در انتشار، انتظار می‌رود که تجمع نقاط سانسور شده در سمت راست نمودار که دارای انحراف استاندارد بیشتری هستند بالاتر باشد. اما در این قسمت نمودار، احتمالاً قرار گیری نقاط سانسور شده یک وضعیت خاص وجود دارد. شانس چاپ مقالاتی که یافته‌های مثبت داشته و اثربخشی را نشان می‌دهند (positive results) بیشتر از نتایج تحقیقاتی است که ارتباط معنی‌داری را پیدا نکرده‌اند. به عبارتی نتایج تحقیقات کوچک معمولاً در شرایط امکان چاپ را پیدا می‌کنند که یک ارتباط معنی‌دار را پیدا نمایند و در صورت عدم مشاهده هر گونه ارتباطی، معمولاً نه محققین و نه مجلات تمایل زیادی به چاپ نتایج ندارند.

پس با این پیش فرضها می‌توان انتظار داشت که در صورت وجود سوگیری در انتشار نتایج نمودار کیفی ترسیم شده در قسمت گشادتر کیف حالت نامتقارن به خود گرفته و در تعداد مقابل ملاحظه‌ای از نقاط سانسور شده احتمالی در یک سمت قسمت گشاد کیف متمرکز شده باشند. برای روشنتر شدن مطلب نمودار شماتیکی رسم شده است.



نمودار:6 شکل شماتیکی از نمودار کیفی که نشان می‌دهد احتمالاً در اثر سوگیری در انتشار نتایج چه نقاطی از این نمودار در جستجوهای سیستماتیک قابل کشف نیستند.

بر اساس توضیحات فوق می‌توان بیان نمود که کاربرد اصلی نمودار کیفی نشان دادن میزان احتمال سوگیری در چاپ (Publication Bias) است. البته باید بیان نمود که از نظر ظاهری این نمودار در مقالات مختلف ممکن است متفاوت باشد و در بعضی مقالات محورهای افقی و عمودی متفاوت باشند ولی مفهوم کلی و اصول ترسیم آنها تقریباً یکسان است.

## ۳-۷. آزمونهای آماری مرتبط به سوگیری در انتشار نتایج

اگر چه این تفسیر از نمودار کیفی برای شناسایی میزان بروز سوگیری در انتشار به نظر ساده و آسان می آید ولی در عمل و تنها با نگاه چشمی به راحتی نمی توان به وجود و یا عدم وجود سوگیری در چاپ اذعان نمود. به همین دلیل سعی شده است همین تفسیر و نتیجه چشمی با آزمون های آماری مناسب نیز تحلیل شوند. در ادامه به دو آزمون اصلی برای بررسی میزان شانس بروز سوگیری در انتشار نتایج اشاره خواهد شد.

### آزمون آماری Begg and Manzumdar

اسم کامل این آزمون Begg and Manzumdar adjusted rank correlation است. در این آزمون همبستگی غیر پارامتریک کندال (Kendall tau) بین شاخص مورد مطالعه و انحراف استاندارد آن در مطالعات مختلف محاسبه می شود.

در صورت معنی دار بودن این ضریب و فاصله گرفتن آن از صفر می توان برداشت نمود که بین مقدار شاخص برآورد شده و اندازه مطالعه ( یعنی حجم نمونه که به صورت معکوس روی انحراف استاندارد تاثیر دارد و با افزایش حجم نمونه مقدار انحراف استاندارد کاهش می یابد )، ارتباط وجود داشته و یا به عبارت دیگر برآورد نقطه ای مطالعات بزرگ و کوچک یکسان نبوده است. انطباق این مفهوم با نمودار کیفی به این شکل است که حالت تقارن نقاط حول خط افقی در قسمت چپ و راست نمودار کیفی یکسان نیست ( رجوع به شکل نمودار 6). چنین یافته ای احتمالاً بدلیل عدم دستیابی به نتایج تعدادی از مطالعات کوچکتر است که اصطلاح

سوگیری در انتشار نتایج در مورد آن بکار گرفته می شود.

نکته مهم در نمودار کیفی و آزمون بگ آن است که بایست لگاریتم شاخصهایی که بصورت نسبت هستند را بکار برد. همچنین معمولاً تمام نرم افزارهای حرفه‌ای قبل از محاسبه ضریب همبستگی، مقدار شاخص را استاندارد می‌نمایند تا جلوی بعضی خطاها گرفته شود. بدین شکل میانگین مقدار شاخص استاندارد شده (Z-Score) برابر صفر خواهد شد.

نهایتاً بر اساس نتایج این آزمون معمولاً ضریب همبستگی و معنی‌داری آماری آن گزارش می‌شود. از آنجایی که معمولاً در متا-آنالیز تعداد مقالات محدود است و ضریب همبستگی به کار گرفته شده در این آزمون غیر پارامتریک است لذا توان (power) این روش عموماً پائین است. بدین جهت در اثر موارد توصیه می‌شود سطح معنی‌داری از ۵٪ افزایش یافته و ۱۰٪ در نظر گرفته شود.

### آزمون آماری Egger regression asymmetry test

روش دیگری که برای بررسی میزان احتمال بروز سوگیری در انتشار نتایج وجود دارد از مدل رگرسیون خطی استفاده می‌کند. در این روش مقدار استاندارد شده شاخص مورد نظر (Z-score) را به عنوان متغیر وابسته در نظر گرفته و مقدار آن را از روی میزان انحراف استاندارد (se) مطالعات پیش بینی می‌کند.

$$Z - \text{score of the effect} = \beta_0 + \beta \times Se$$

در صورت عدم وجود سوگیری در چاپ (Publication Bias) مقدار  $\beta_0$  به عنوان عدد ثابت معادله رگرسیون باید به صفر نزدیک باشد چرا که  $\beta_0$  عبارت است از میانگین شاخص استاندارد شده برای زمانی که انحراف استاندارد به سمت صفر میل نماید و از آنجایی که میانگین شاخص استاندارد شده (Z-score) برابر صفر است، لذا انتظار می‌رود در صورت عدم وجود سوگیری در انتشار نتایج، مقدار  $\beta_0$  صفر باشد.

مطالعات مختلف نشان داده اند که میزان قدرت (Power) آزمون Egger بیشتر از آزمون Begg است و سریعتر احتمال وجود سوگیری در چاپ را نمایش می‌دهد. به همین دلیل در حجم نمونه‌های کم، ترجیح داده می‌شود که از روش Egger استفاده شود.

یک سؤال آسان ولی دشوار!

آیا می‌توانید حدس بزنید دلیل اصلی توان بیشتر آزمون Egger چیست؟

پاسخ دادن به این سؤال برای کسانی که با آزمونهای پارامتریک و غیر پارامتریک آشنا هستند آسانتر خواهد بود. البته پاسخ این سوال در متن آزمونها نیز نهفته است. اگر پاسخ سوال را نبافتید، یکبار دیگر دو سه صفحه قبل را مطالعه نمایید.

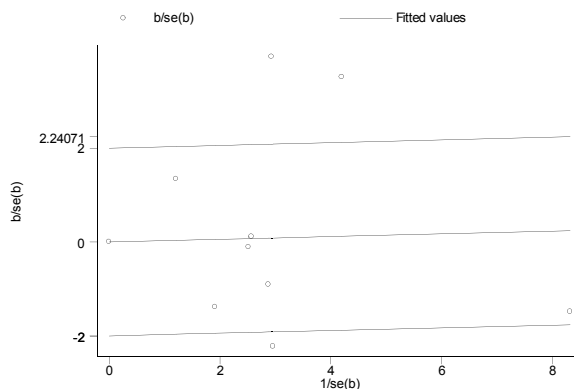
## ۷-۴. نمودار Galbraith

این نمودار برای نشان دادن میزان احتمالی عدم تجانس بین مطالعات استفاده می‌شود و به صورت گرافیکی نشان می‌دهد که احتمالاً نتایج کدام مطالعه نسبت به بقیه عدم همخوانی دارد.

همانگونه که در قسمت ۶-۹ شرح داده شده است، آزمون آماری مورد استفاده برای بررسی عدم تجانس معمولاً دارای توان (Power) آماری پائین در کشف میزان عدم تجانس است. چرا که تعداد مطالعات وارد شده در متاآنالیز کم بوده و لذا حجم نمونه این آزمون آماری پائین است. نمودار Galbraith که در سال ۱۹۸۸ شرح داده شد یک روش نسبتاً آسان و دقیق برای کشف عدم تجانس است. در این نمودار مقدار استاندارد شاخص مورد نظر (Z-score) بر روی محور Y و مقدار وزن مطالعات یعنی معکوس واریانس بر روی محور X نمایش داده می‌شود. لازم به ذکر است که همانند نمودار های قبلی، در صورتی که شاخص مورد نظر از نوع نسبت باشد (RR, OR) و یا HR) ابتدا باید از آن لگاریتم گرفت و سپس استاندارد کرد. در مرحله بعد خط رگرسیونی رسم می‌شود که محور Y را در مبدأ یعنی عدد صفر قطع نماید و ضریب زاویه آن (یعنی شیب خط) برابر مقدار پارامتر برآورد شده در مدل فیکس است. سپس از نقاط ۲- و ۲+ بر روی محور Y دو قطع موازی با خط رگرسیون رسم می‌شود تا دامنه اطمینان ۹۵٪ را نمایش دهد. حال انتظار می‌رود که در صورت نبودن عدم تجانس (Heterogeneity) در نتایج، تمام نقاط مربوط به مطالعات در بین دو خط ۲+ و ۲- یعنی در دامنه اطمینان ۹۵٪ قرار گیرند. در صورتی که بعضی نقاط خارج از این دامنه قرار گیرد

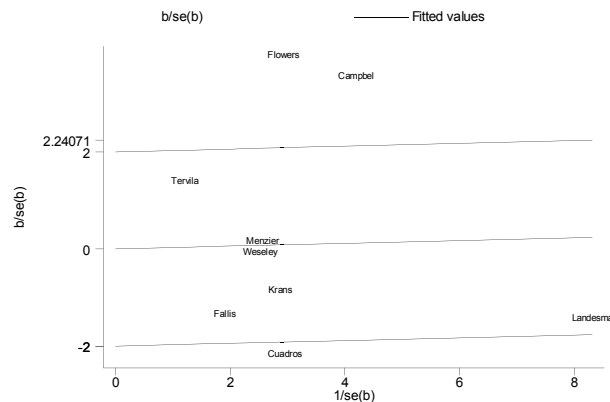


نشان دهنده عدم تجانس نتایج آن مطالعات با بقیه است.



نمودار ۷: نمودار Galbraith برای نشان دادن میزان عدم تجانس بین نتایج مطالعات وارد شده در متا-آنالیز مربوط به نسبت خطر مصرف دیورتیک و بروز پره اکلامپسی. همانگونه که ملاحظه می‌فرمایید یک نقطه در پائین و دو نقطه در بالای خطوط دامنه اطمینان ۹۵٪ قرار گرفته‌اند و بنابراین بیشترین تاثیر را در بروز عدم تجانس بین نتایج مطالعات دارند.

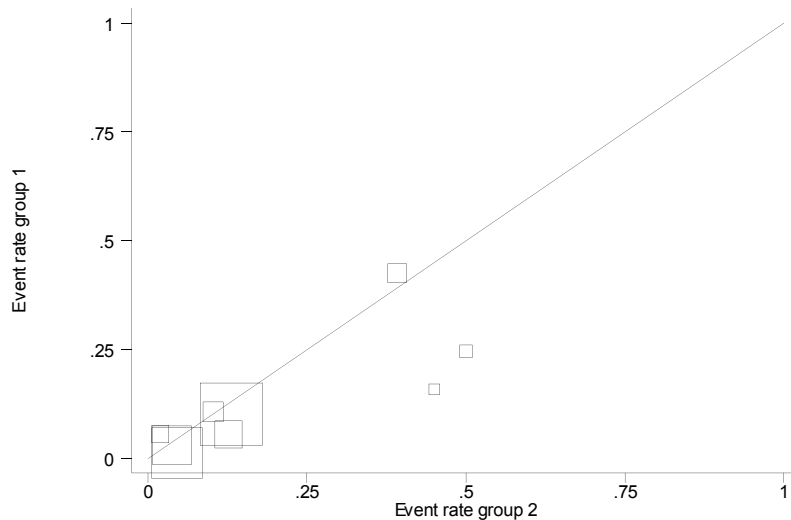
در این روش حتی می‌توان دریافت که نقاط مربوط به کدام مطالعات خارج دامنه مورد نظر قرار دارند که باعث عدم تجانس می‌شود و سپس با شناسایی این مطالعات می‌توان با دقت بیشتر علت عدم تجانس این مطالعات را با بقیه کشف نمود.



نمودار ۸: نمودار Galbraith همانند شکل قبلی، فقط در این شکل نام مطالعات بر اساس اسم نویسنده اول تعیین شده است و لذا می‌توان راحت‌تر دریافت که کدام مطالعات باعث عدم تجانس نتایج شده اند

### ۷-۵. نمودار L'Abbe

این نمودار فقط زمانی کاربرد دارد که متغیر مورد اندازه‌گیری در گروه‌های مورد و شاهد یک متغیر دو حالتی مانند بهبود و عدم بهبود و یا مرگ و زندگی باشد. در این نمودار فراوانی (درصد و یا میزان) وقوع پیامد در گروه‌های مورد و شاهد در محورهای افقی و عمودی نمایش داده می‌شود و سپس مقادیر متناظر هر مطالعه در نمودار نشان داده می‌شود. معمولاً نقاط نشان دهنده اطلاعات هر مطالعه به صورت دایره و یا مربعهایی هستند که مساحت آنها متناسب با وزن آنها می‌باشد. همچنین در این شکل معمولاً خط متناظر با فرضیه صفر (عدم اختلاف میزان وقوع پیامد در گروه‌های مورد و شاهد) نیز رسم می‌شود.



نمودار ۹: نمودار L'Abbe؛ این نمودار درصد وقوع پیامد در گروههای مورد و شاهد را به تفکیک مطالعه نشان می‌دهد و مساحت هر مربع متناسب با وزن مطالعه مورد نظر است.

با مشاهد اطلاعات موجود در نمودار L'Abbe می‌توان دریافت که فراوانی وقوع پیامد مورد مطالعه در کدام مطالعات بیشتر است و آیا فراوانی وقوع پیامد به نتیجه‌گیری نهایی مطالعه مرتبط است یا خیر؟

نکته مهمی که در رسم این نمودار به آن تاکید می‌شود آن است که بهتر است در این نمودار خط رگرسیون که نشان دهنده برآیند مقادیر اثر مداخله (به عبارتی نسبت وقوع پیامد در گروه مورد و شاهد است) را رسم نماییم. دلیلی اصلی آن این است که در رگرسیون خطی، وزن مطالعات و همچنین انحراف استاندارد مقادیر بدست آمده در گروه مورد و شاهد را در نظر نمی‌گیرد و به همین دلیل نتایج آن ممکن است منحرف کننده بوده و با نتایج متا-آنالیز متفاوت باشد. البته اگر شما علاقمند به رسم خط رگرسیون باشید می‌توانید با در نظر

گرفتن موارد فوق، خط رگرسیون را نیز به نمودار خود اضافه نمایید.

نمودار L'Abbe نیز همانند نمودار کیفی تا حدودی می-تواند احتمال بروز سوگیری در انتشار نتایج را نشان دهد. بدین معنی که اگر نقاط متناظر با مطالعات کوچکتر که با مربعهای کوچکتر در شکل نمایش داده می-شوند بیش از مربعهای بزرگ در یک طرف خط متمرکز باشند احتمال سوگیری در انتشار را مطرح می-کند. البته پر واضح است که تفسیر و نتیجه‌گیری بر اساس نمودار کیفی آسانتر است و به همین جهت میزان استفاده از نمودار L'Abbe بسیار محدودتر است.

---

## ۸. موضوعات ویژه در متاآنالیز

---

در فصول قبل، کلیات متاآنالیز و مبانی آماری آن شرح داده شد و سعی گردید با کمترین گرایش به فرمول های آماری مربوط، مفاهیم تا حد امکان آسان و روان شرح داده شود. ولی علاوه بر موضوعات مطرح شده مذکور، در حال حاضر مفاهیم جدیدتری نیز در مباحث متاآنالیز وجود دارد که دانستن این مفاهیم برای محققین و دانشجویان بسیار مفید است. در این فصل به مهمترین موضوعات در این زمینه اشاره خواهد شد.

### ۸-۱. متاآنالیز تجمعی Cumulative Meta-Analysis

در متاآنالیز معمولی سعی می‌شود اطلاعات مطالعات با یکدیگر و در یک مرحله ترکیب شود و برآیند آنها تعیین گردد. در متاآنالیز معمولی ترتیب مطالعات اهمیتی ندارد و ترکیب نهایی آنها با هر ترتیبی یکسان است. ولی در متاآنالیز تجمعی ترتیب مطالعات اهمیت پیدا می‌کند و معمولاً سال چاپ، عامل اصلی مرتب نمودن اطلاعات در نظر گرفته می‌شود.

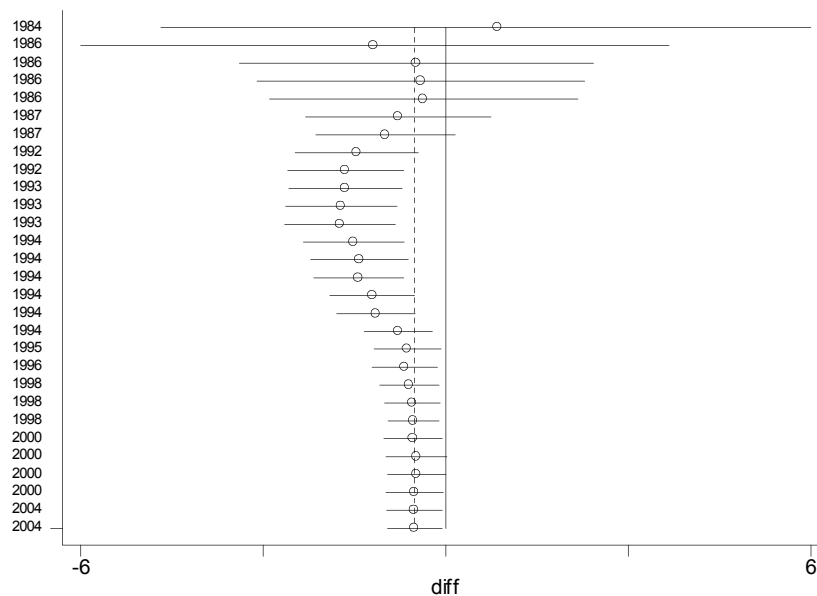
بدین شکل که به ترتیب اطلاعات قدیمی‌ترین منبع به عنوان پایه در نظر گرفته می‌شود و سپس اطلاعات

مطالعات بعدی به آن اضافه می‌شود و به همین شکل اطلاعات مطالعات جدیدتر به مطالعات قبلی اضافه می‌شوند. طبیعتاً با اضافه شدن مطالعات، به تدریج دقت متاآنالیز افزایش یافته و در نتیجه دامنه اطمینان محاسبه شده کوچکتر و کوچکتر می‌شود.

در مراحل اول، اضافه شدن نتایج یک مطالعه ممکن است برآورد نقطه‌ای را به صورت قابل ملاحظه‌ای جابجا کند چرا که وزن مطالعه جدید در مقابل مطالعات محدود قبلی قابل ملاحظه است ولی در مراحل آخر، برآورد بدست آمده حاصل ترکیب اطلاعات تعداد بیشتری مطالعه است و لذا وزن زیاد خواهد داشت و اضافه شدن مطالعات جدید کمتر می‌تواند تأثیر قابل ملاحظه‌ای در جابجایی برآورد نقطه‌ای اعمال نماید.

کاربرد این دسته از مطالعات متاآنالیز تجمعی، زمانی است که بخواهیم مشخص‌نمائیم که احتمالاً از چه زمانی مستندات کافی برای رد یا قبول مثلاً یک روش درمانی وجود داشته است.

به عنوان مثال در نمودار زیر میزان نفوذ رنگ بعد از ترمیم ریشه دندان در دو روش مختلف مقایسه شده است. در این تحقیق شاخص مورد بررسی تفاوت میزان نفوذ رنگ در روش اول با روش دوم بوده است. همانگونه که مشاهده می‌کنید ترکیب اطلاعات تحقیقات تا سال ۱۹۹۲ اختلاف معنی‌داری را بین دو روش نشان نمی‌دهند ولی از آن سال به بعد ترکیب اطلاعات مطالعات مختلف نتیجه‌گیری نهایی را تغییر نداده است و نشان می‌دهد که میزان نفوذ رنگ در روش اول کمتر از روش دوم است.



### ۲-۸. متا-آنالیز P-values

در سال ۱۹۳۲، فیشر در خصوص ترکیب p-value ها مطالبی را نگاشت که به نظر جونز (۱۹۹۵) منشا متا-آنالیز بوده است. به هر شکل، این ایده وجود دارد که با ترکیب نمودن p-value ها بتوان تصمیم کلی در خصوص معنی-داری نتایج کل مطالعات ارائه نمود. البته باید توجه داشت که این ترکیب ممکن است به شدت گمراه کننده باشد. چرا که معمولاً در مقالات p-value دودامنه بیان می‌گردد و نشان می‌دهد که مثلاً آیا میانگین در دو گروه تفاوت معنی‌دار دارند یا خیر؟ در صورت معنی‌دار بودن p-value نمی‌توان گفت که میانگین در گروه اول بیشتر از دوم بوده است و یا بالعکس. در نتیجه ممکن است در بعضی مطالعات p-value معنی‌دار به نفع بیشتر بودن

میانگین در گروه اول و در بعضی مطالعات برعکس باشد. در نتیجه طبیعی است که تصور شود که این مطالعات باید اثرات یکدیگر را خنثی نمایند. ولی در ترکیب  $p$ -value عملاً این نکته قابل اعمال نیست و به عبارتی فقط قدر مطلق تفاوتها را بیان میکند و نه جهت تفاوتها را.

از طرف دیگر ممکن است  $p$ -value را به روشهای مختلف و با آزمونهای آماری متفاوتی محاسبه شده باشند که این نکته نیز باید مد نظر قرار گیرد. به عبارتی به راحتی نمیتوان فقط و فقط  $p$ -value ها را از مقالات استخراج و در فرمولهای مربوط وارد و نتیجه گیری نمود.

آخرین نکته قابل بحث آن این است که متا-آنالیز  $p$ -value معنی داری را بررسی میکند و شدت ارتباط و یا اختلاف را نمیسنجد. به همین دلیل نمیتوان با دیدن نتایج این نوع متا-آنالیز در خصوص **force of association** بحث نمود. همچنین این نوع تحلیل اثرات عدم تجانس بین مطالعات (**heterogeneity**) را در نظر نمیگیرد.

با توجه به توضیحات فوق میتوان بیان نمود که مهمترین کاربرد متا-آنالیز  $p$ -value در ترکیب نتایج مطالعاتی بکار می رود که برای تحلیل اطلاعات خود از روشهای آماری غیرپارامتریک استفاده نموده اند. همانگونه که احتمالاً مستحضرید مهمترین کاربرد آزمونهای آماری غیرپارامتریک زمانی است که متغیرهای مورد تحلیل از نوع اسمی و عمدتاً رتبه ای هستند. برای ترکیب  $p$ -value سه فرمول وجود دارد که در ادامه به آنها اشاره و محاسن و معایب هر یک بیان خواهد شد.

### فرمول فیشر

فرمول این تحلیل به صورت زیر است



$$U = -2 \sum_{j=1}^k \ln(p_j)$$

در این فرمول،  $k$  نشان دهنده تعداد **p-value** هایی است که با هم ترکیب می‌شوند. به عبارتی برای بدست آمدن  $U$  باید لگایتم طبیعی مقادیر **p-value** ها را با یکدیگر جمع نمود و بعد در منهای دو ضرب کرد. مقدار  $U$  از توزیع کای-دو پیروی می‌کند و درجه آزادی آن برابر  $2k$  می‌باشد. لذا بعد از بدست آمدن مقدار  $U$  باید مقدار **p-value** کل را از جدول کای-دو استخراج نمود. لازم به ذکر است که **p-value** های بکار رفته در این فرمول باید یک طرفه باشند و در صورتی که در مقالات به صورت دو طرفه بیان شده‌اند باید قبل از استفاده بر دو تقسیم شوند تا به صورت یک طرفه در آیند. این فرمول به نظر بسیاری از متخصصان در صورتی که تعداد مطالعات زیاد باشد، حساسیت بالا دارد و بیش از حد نتیجه نهایی را معنی‌دار بیان می‌کند.

### روش Edgington

این دانشمند دو فرمول مجزا برای ترکیب **p-value** ها ارائه نموده است که به ترتیب عبارتند از:

$$p = \frac{(\sum_{j=1}^k p_j)^k}{k!}$$

این روش به خصوص برای ترکیب **p-value** ها در چند مطالعه محدود کاربرد دارد و نتیجه مطلوب ارائه می‌دهد ولی اگر تعداد مقالات ترکیبی زیاد شود ممکن است جمع کل از یک بیشتر شود. برای رفع این ایراد فرمول دیگری ارائه شده که به شکل زیر است:

$$\bar{p} = \frac{\sum_{j=1}^k p_j}{k} \quad U = (0.5 - \bar{p})\sqrt{12}$$

در این فرمول ابتدا باید مقدار  $p$  را حساب نمود و بعد مقدار  $U$  را بدست آورد.  $U$  از توزیع نرمال پیروی میکند و میتوان در جدول توزیع نرمال  $p$ -value نهایی را بدست آورد. این فرمول در محاسبه  $p$ -value از ترکیب حداقل ۴ مطالعه کاربرد دارد.

خوشبختانه تمام این تحلیلهای توسط برنامه `stata` انجام میشود و توضیحات مربوطه در فصل بعد آورده شده است.

### ۳-۸. ترکیب اطلاعات تک تک نمونه‌ها مطالعات (IPD)

همانگونه که تا کنون شرح داده شده است در روشهای معمول متا-آنالیز، اطلاعات نهایی مطالعات که معمولاً در مقالات و یا گزارشهای تحقیقات وجود دارد با یکدیگر ترکیب و نتیجه نهایی گرفته میشود. چرا که در اکثر اوقات امکان دسترسی به ریز اطلاعات و بانک اطلاعات خام مطالعات مقدور نیست. به همین دلیل نیز بعضی خطاها مانند خطای `aggregating bias` ممکن است که نتیجه را مخدوش نماید. همانگونه که قبلاً نیز توضیح داده شده این خطا عبارت است از تعمیم نتایج گروه به افراد؛ به عبارتی بعضی یافته‌های گروهی برای افراد قابل تعمیم نیست.

برای حل این مشکل، روشهای جدیدی بیان شده است که در آنها به جای ترکیب اطلاعات نهایی و نتیجه‌های مطالعات، اطلاعات تک‌تک نمونه‌ها با یکدیگر ترکیب می-

شوند. این روش اصطلاحاً روش **Individual Patient Data (IPD)** نامیده می‌شود. به عبارتی این روش را می‌توان به نوعی آنالیز مجدد اطلاعات خام مطالعات دانست.

البته انجام **IPD** مزایای دیگری نیز دارد که مهمترین آنها به شرح زیر است:

۱. می‌توان فرضیه‌هایی که در مطالعات قبلی بررسی نشده‌بوده‌اند را نیز آزمود. به عنوان مثال می‌توان اثرات فاکتورهای زمینه‌ای را نیز مطالعه نمود.

۲. می‌توان خطاهای مخدوش‌کننده را به دقت و با یک روش یکسان در کلیه مطالعات بررسی و برطرف نمود.

۳. با دقت بیشتری می‌توان مقادیر نامعلوم برای بعضی از نمونه‌ها را مدیریت نمود.

۴. توان نهایی آزمون آماری بالاتر خواهد بود.

۵. و نهایتاً باعث افزایش سطح همکاری‌های بین‌المللی خواهد شد.

پر واضح است که بزرگترین مشکل چنین تحلیلهایی دستیابی به اطلاعات تک‌تک نمونه‌ها است. همانگونه که در فصلهای قبل به تفصیل شرح داده شده است، در حال حاضر امکان دستیابی به اطلاعات نهایی مقالات و تحقیقات به راحتی مقدور نیست و سوگیری در انتشار نیز منتج از همین محدودیت است، حال تصور نمایید که شما علاقمند به انجام مطالعه مروری ساختار یافته و ترکیب نتایج به شیوه **IPD** باشید. تقریباً شانس شما برای یافتن ریز اطلاعات نمونه‌های وارد شده در

مطالعات بسیار پائین‌تر خواهد بود. به همین لحاظ با وجود مزیت‌های قابل توجه این روش، کاربرد آن هنوز محدود است.

برای حل مشکل دستیابی به ریز اطلاعات، مراکزی در سطح دنیا به جمع‌آوری بانک‌های اطلاعات تک‌تک مطالعات کارآزمایی بالینی می‌پردازند. با این امید که با گذشت زمان این بانک‌ها به محل غنی و مناسبی برای انجام تحلیل‌های IPD تبدیل شوند. البته با توجه به اهمیت مطالعات کارآزمایی‌های بالینی در متا-آنالیز، این بانک‌ها تنها به جمع‌آوری اطلاعات مربوط به این دسته از مطالعات می‌پردازند.

نکته مهم دیگری باید به آن توجه داشت آن است که در صورت دستیابی به کل اطلاعات مطالعات مرتبط، انجام تحلیل IPD باید با اصول و متدهای خاصی صورت پذیرد و نمی‌توان به راحتی تمام اطلاعات را در یک بانک اطلاعات کلی وارد نمود و با آزمون‌های آماری معمول به تحلیل نتایج پرداخت. مهمترین نکته در تحلیل چنین مطالعاتی، در نظر گرفتن *Intra-cluster correlation* است. به عبارتی، باید از مدل‌های آماری استفاده نمود که ارتباطات درون گروهی را به عنوان یک فاکتور تعدیل کننده واریانس در مدل وارد می‌کنند. لذا به نوعی می‌توان بیان نمود که چنین تحلیل‌هایی شبیه به تحلیل اطلاعات مطالعات چند مرکزی (*multi-center trials*) است.

### ۴-۸. تحلیل حساسیت (sensitivity analysis)

یکی از بحث‌های اصلی در متا-آنالیز بررسی میزان تاثیر تک تک مطالعات در نتیجه نهایی است. در صورتی که نتایج یک یا چند مطالعه به صورت قابل توجه باعث

تغییر نتیجه‌گیری نهایی شود، اعتماد به پاسخ نهایی کمتر خواهد شد. معمولاً زمانی نتیجه نهایی از اطلاعات یک مطالعه به شدت متاثر می‌شود که نتیجه آن مطالعه از جمع نتایج بقیه مطالعات فاصله گرفته‌باشد و به عبارتی پاسخ اغراق‌آمیز داشته‌باشد و همچنین نسبت حجم نمونه (و یا به عبارت دقیق‌تر وزن مطالعه) مطالعه مذکور به وزن سایر مطالعات قابل توجه باشد.

معمولاً در متا-آنالیز برای نشان دادن اینکه تاثیر هر مطالعه بر روی نتیجه نهایی چقدر است، روشهای مختلفی وجود دارد. یکی از ساده‌ترین روشها آن است که به ترتیب تک‌تک مطالعات را از تحلیل خارج نمود و نتیجه بدست آمده با و بدون وجود مطالعه مذکور را با هم مقایسه نمود. در صورتی که بین این دو تفاوت قابل ملاحظه باشد، و به عبارتی، حساسیت نتیجه نهایی به وارد شدن نتایج آن مطالعه بالا باشد، اصطلاحاً عنوان می‌شود که میزان ثبات (robustness) مدل پایین است.

طبیعاً با افزایش تعداد مطالعات حتی در صورت وجود پاسخهای متنوع و وجود عدم تجانس بسیار زیاد بین مطالعات، میزان ثبات افزایش می‌یابد چرا که نسبت وزن تک‌تک مطالعات به جمع وزن سایر مطالعات کاهش خواهد یافت.

البته در روشهای پیچیده‌تر سعی خواهد شد تا به جای حذف صرفاً تک‌تک مطالعه در هر مرحله ترکیبی از مطالعات حذف شود. مثلاً به صورت تصادفی ۲ تا ۲ از نتایج مقالات حذف و متا-آنالیز بر روی بقیه مطالعات انجام شود. پر واضح است که در این صورت ترکیبات دوتایی بسیاری به وجود خواهد آمد. به عنوان مثال اگر ۵ مطالعه داشته‌باشیم ترکیب دوتایی می‌تواند ترکیب مطالعه یک با دو، یک با سه، ..... باشد.

مشخص نمودن دستی این ترکیبات و خارج نمودن آنها در هر مرحله و مقایسه نتایج بدون آن ترکیب دوتایی با نتیجه کل، کار بسیار زمانبری است و بدون برنامه‌های کامپیوتری آماری انجام آن تقریباً غیر ممکن است. خوشبختانه برنامه **stata** این توانایی را دارد و با نوشتن چند خط برنامه می‌توان به نتیجه مورد نیاز دست یافت که در فصل بعدی به بعضی از این توانمندیهای اشاره خواهد شد.

البته مدل‌های پیچیده‌تر آماری نیز وجود دارد که سعی می‌کند میزان تاثیر مطالعاتی که در يك سوي شاخص نهایی قرار دارند و سنجش میزان سوگیری در انتشار نتایج، شاخص نهایی را تصحیح نماید و به عبارتی میزان تعدیل شده‌ای را ارائه کند که تا حد امکان اثر سوگیری در انتشار نتایج را خنثی نماید. در این شیوه، ابتدا میزان احتمال بروز سوگیری در انتشار با استفاده از روشهایی که قبلاً ذکر شد تعیین می‌گردد (روش **Begg** و یا **Egger**). سپس با استفاده از مدل‌های شناخته‌شده‌ای مانند روشهای رگرسیون خطی و یا غیرپارامتریک سعی می‌شود نقاط متناظری برای نقاط ایجاد کننده سوگیری در انتشار پیش‌بینی شده و سپس مدل متا-آنالیزی با وارد نمودن نقاط پیش‌بینی شده اجرا می‌شود.

یکی از شناخته‌شده‌ترین مدل برای انجام این گونه تحلیلها که در برنامه **stata** نیز از آن استفاده می‌شود به نام **Duval and Tweedie nonparametric "trim and fill"**

معروف است. توضیحات بیشتر در خصوص شیوه اجرای این دستور در برنامه **stata** در فصل بعدی آورده خواهد شد.

## ۸-۵. تاثیر نمره کیفیت مطالعات در وزن داده شده در متا-آنالیز

همانگونه که در فصلهای قبل شرح داده شد، به صورت کلی وزن داده شده به مطالعات معکوس انحراف استاندارد مطالعات است؛ و انحراف استاندارد خود متأثر از حجم نمونه است. بنابراین، مهمترین فاکتور تاثیرگذار بر وزن مطالعات در روشهای کلاسیک متا-آنالیزی اندازه مطالعه است و چه مطالعه بزرگتر باشد و حجم نمونه بیشتری داشته باشد وزن بیشتری را نیز به خود اختصاص می دهد. البته در فصلهای قبل مفصل بیان شد که تاثیر حجم نمونه بر وزنها در مدل های ثابت بسیار بیشتر از مدل های تصادفی می باشد.

اما سوالی که مطرح است این است که آیا می توان وزن مطالعات را بر اساس کیفیت مطالعات نیز تعدیل نمود؟ به عبارتی به نتایج مطالعه ای که با دقت بیشتر انجام شده وزن بیشتری تعلق بگیرد.

بر این این تعدیل باید نمره ای را برای کیفیت مطالعات محاسبه نمود. در حال حاضر فرمولهای متعددی در خصوص محاسبه چنین امتیازی پیشنهاد می شود و در سایت ککرین توضیحات مفصلی در این خصوص وجود دارد که از محدوده این کتاب خارج است ولی فقط برای نمونه به بعضی از گزینه ها اشاره می شود:

۱. آیا تحقیق توسط یک سازمان و یا دانشگاه معتبر از نظر علمی و یا مالی پشتیبانی شده است؟

۲. آیا محقق اصلی تحقیق یک فرد کارپشته و محقق برجسته بوده است؟

۳. آیا مقاله در يك مجله معتبر و با impact factor بالا به چاپ رسیده است؟

۴. آیا در تیم تحقیق افراد متخصص با گرایشهای مختلف در زمینه رشته علمی مربوطه، متدولوژی تحقیق، متخصص آمار و اپیدمیولوژی وجود داشته است؟

معمولاً برای محاسبه امتیاز کیفیت مطالعات، نمره ای را به هر گزینه می‌دهند و سپس جمع ساده و یا جمع وزن- داده شده امتیازات گزینه‌ها را محاسبه و با هم جمع می‌کنند.

البته به صورت يك اصل كلي و مورد توافق تقریباً تمامی صاحب‌نظران، اگر مطالعه ای دقت کافی ندارد نباید اصلاً در متا-آنالیز وارد شود. لذا اگر صحبت در خصوص وزن دادن به مطالعات بر اساس کیفیت مطالعات می‌شود، منظور مطالعاتی هستند که حداقل کیفیت را دارا هستند و عملاً توضیحات بعدی این امکان را فراهم نمی‌کند که بتوانیم از نتایج مطالعات پر خطا و دارای ایرادات اصلی متدولوژی استفاده نماییم.

در برخورد با امتیاز کیفیت مطالعاتی که حداقلها را دارا هستند چند تکنیک قابل ارایه است. در یکی از این تکنیکها سعی می‌شود نمره تعیین شده برای کیفیت مطالعات را به عنوان يك متغیر مستقل در مدل متا-رگرسیون وارد نمود. در این مدل، سعی می‌شود نه تنها اثر کیفیت مطالعه ارزیابی گردد بلکه تاثیر آن حذف و نتیجه نهایی تصحیح شده بیان خواهد شد. این روش اگرچه مزایای بسیار دارد ولی يك ایراد اصلی بر آن وارد است و آن این است که فرض می‌کند امتیاز کیفیت داده به مطالعات يك امتیاز عددی واقعی است و این در



حالی است که تقریباً فرض عددی بودن این امتیاز، خالی از اشکال نیست.

در روش دیگر، سعی می‌شود که مطالعات را بر اساس امتیاز کیفیتشان، به چند زیر گروه تقسیم نمود و سپس در هر زیر گروه تحلیلهای مربوطه را صورت جداگانه در زیر گروهها انجام داد. این روش اگرچه از نظر متدلوژی قابل دفاعترین روش است ولی يك ایراد اصلی دارد و آن محدودیت حجم نمونه است. معمولاً در متا-آنالیز تعداد مقالات و مطالعات وارد شده محدود است و تقسیم آنها به چند زیر گروه مشکل را چند برابر می‌کند و تعداد مطالعات در هر زیر گروه را به شدت کاهش می‌دهد.

تکنیک آخر آن است که وزن مطالعات را با توجه به امتیاز کیفیت آنها تعدیل نمود. در این روشها وزن مطالعه بر اساس فرمول زیر محاسبه می‌گردد:

$$w = \frac{\text{Quality Score}}{se}$$

به عبارتی وزن مطالعه نسبت مستقیم با کیفیت مطالعه و نسبت معکوس با انحراف استاندارد مطالعات خواهد داشت.

## فصل نهم

---

### ۹. دستورات برنامه **stata** برای انجام متا-آنالیز

---

برنامه **stata** به عنوان یکی از کاملترین و متداولترین نرم افزارهای آماری مورد استفاده در آمار حیاتی و اپیدمیولوژی مطرح است و توانمندیهای این برنامه به شدت مورد علاقه متخصصانی است که برای کارهای حرفه‌ای خود نیاز به یک نرم افزار جامع دارند.

یکی از قابلیت‌های بسیار ارزشمند این برنامه توانمندیهای این برنامه در انجام متا-آنالیزهای ساده و پیچیده است. از آنجایی که تقریباً تمامی مدل‌های رایج متا-آنالیز در این برنامه قابل اجراست به همین جهت در این کتاب سعی می‌شود تا دستورات مربوطه به صورت مختصر شرح داده شود تا علاقمندان بتوانند بحث‌های تئوری فصول قبل را عملاً بکار ببرند و در صورت انجام متا-آنالیز، شخصاً با استفاده از توانمندیهای این برنامه بهره ببرند.

طبیعی است که برای انجام متا-آنالیز شما باید با نحوه کار با برنامه آشنا باشید و بتوانید کارهای ساده آماری را در این برنامه انجام دهید. از آنجایی که این کتاب فقط برای آشنا نمودن قابلیت‌های متا-آنالیزی در این برنامه است لذا از کاربران محترم باید قبل از مطالعه و اجرای دستورات این بخش یا در کارگاه‌های آموزشی **stata** شرکت نماید و یا با مراجعه به

منابع موجود سعی نماید اطلاعات کافی را در بکارگیری عمومی این برنامه کسب نماید. لازم به ذکر است که در اینترنت منابع آموزشی بسیار زیادی برای فراگیری کار با این نرم افزار وجود دارد. برای کسب اطلاعات بیشتر توصیه می‌شود کاربران محترم به آدرس اینترنتی این برنامه ([www.stata.com](http://www.stata.com)) مراجعه نمایند.

در انتها مجدد بیان می‌شود که این فصل کتاب به عنوان یک جزوه آموزشی برای استفاده در کارگاه‌های متا-آنالیز تنظیم شده است و استفاده از آن به عنوان خودآموز به خصوص در صورت عدم آشنایی با برنامه **stata** بسیار دشوار است و اگر شما قصد دارید با مطالعه این فصل به توانمندی کافی برای استفاده از **stata** در انجام متا-آنالیز برسید باید وقت زیادتری صرف نموده و همگام با مطالعه این فصل با برنامه نیز کار نمایید.

### ۹-۱. نحوه اجرای دستورات متا-آنالیز

در این قسمت کاربران محترم باید به دو نکته توجه نمایند. اول آنکه دستورات متا-آنالیز جز اصلی برنامه **stata** نیستند و باید از وب سایت برنامه این دستورات دریافت و نصب شوند. نکته دوم آن است که با وجود وارد شدن منوهای اجرای دستورات در برنامه **stata** از ویرایش هشتم به بعد، متأسفانه دستورات متا-آنالیز باید در برنامه تایپ شوند و در برنامه منوی خودکاري برای انتخاب دستورات وجود ندارد که این نکته باز خودآموزی این فصل را دشوارتر می‌کند. البته در ادامه با دستورات مرتبط متا-آنالیز در برنامه به تفکیک بیان و شیوه نگارش آنها آورده می‌شود تا به شما در آینده کمک نماید. لازم به ذکر است که برنامه **stata** این قابلیت را دارد که منوهای جدیدی از جمله منو

مربوط به متا-آنالیز را در آن وارد نمود ولی برای مختصر نمودن بحث و خارج نشدن از حیطه بحث از ذکر آن خودداری می‌شود. علاقمندان برای آشنایی با این موضوع می‌توانند در قسمت جستجوی برنامه عبارت `meta_dialog` را تایپ و از برنامه بخواهند که تمام منابع را برای یافتن این عبارت جستجو کند.

برای نصب دستورات دو راه وجود دارد. راه حل اول آن است که در برنامه `stata` در حالی که به اینترنت وصل هستید دستور `update all` را اجرا نمایید و به برنامه اجازه دهید که به صورت اتوماتیک تمام برنامه‌ها و دستورات اضافه را بخ صورت خودکار بر روی سیستم شما نصب کند. طبیعی است که در این حالت باید مطمئن باشید که برنامه شما از نوع `original` باشد و شما به صورت رسمی مجوز دستیابی به این دستورات را داشته-باشید.

راه حل دوم شاید آسانتر باشید. برای اینکه باید در وب سایت برنامه `stata` و کمک گرفتن از جستجوی آن فایل‌های دستورات را یک به یک جستجو نموده و بر روی دستگاه خود نصب کنید. از آنجایی که تعداد دستورات مورد استفاده برای متا-آنالیز محدود است لذا به نظر می‌رسد این شیوه زیاد وقت گیر نباشد. معمولاً برای هر دستور باید به دنبال دو فایل باشید. یکی از این فایل‌ها فایل اجرایی دستور مورد نظر است و فایل دیگر فایل راهنما است. به عنوان مثال برای اجرای دستور `meta` که یکی از کلیدی‌تری دستور است نام کامل فایل‌های مورد نیاز عبارتند از:

`meta.ado` فایل اجرایی دستور است و با نصب این فایل برنامه می‌تواند کار خود را آغاز نماید.

`meta.hlp` فایل راهنما است و در صورت نصب، با نوشتن دستور زیر، برنامه راهنمای نسبتاً کافی و کاملی را در اختیار شما قرار می‌دهد و شما می‌توانید در حین کار از آن استفاده نمایید. نحوه درخواست راهنما برای هر دستور `stata` عبارت است از

`help <command>` for example `help meta`

بعد از دریافت فایل دستورات و راهنمای آنها، باید این فایلها را در یک شاخه‌ای بنام `ADO` که به صورت خودکار در هنگام نصب برنامه `stata` در ریشه `c:` ساخته شده است کپی شوند. توجه داشته باشید که این فایلها نیاز به نصب ندارند و فقط باید کپی شوند.

## ۹-۲. دستورات اصلی متا-آنالیز

برای آسانتر شدن نحوه کار با برنامه، ابتدا ساختار فایلها را با هم مرور می‌کنیم که اطلاعات مربوط به مطالعات اثر بخشی مصرف دیورتیکها بر بروز پره-اکلامپسی را شرح می‌دهد. متغیرها و سطرهای این فایل عبارت هستند از:

trial	p_t	p_c	np_t	np_c	OR	logOR	VlogOR	SElogOR
Menzier	14	24	43	24	0.33	-1.12	0.18	0.42
Fallis	6	18	32	22	0.23	-1.47	0.30	0.55
Cuadros	12	35	999	725	0.25	-1.39	0.11	0.34
Landesman	138	175	1232	1161	0.74	-0.30	0.01	0.12
Krans	15	20	491	504	0.77	-0.26	0.12	0.35
Tervila	6	2	102	101	2.97	1.09	0.69	0.83
Campbel	65	40	88	62	1.14	0.14	0.07	0.26

ستونهای `p_t`, `p_c`, `np_t`, `np_c` به ترتیب عبارت هستند از تعداد افراد غیر مبتلا به پره‌اکلامپسی در گروه کنترل و مورد، و همچنین افراد مبتلا به پره‌اکلامپسی در این

دو گروه. در ستونهای بعدی به ترتیب نسبت خطر، لگاریتم طبیعی آن و واریانس و انحراف معیار نسبت خطر محاسبه شده در مطالعات آورده شده است. هر سطر آن جدول نیز اطلاعات یک مطالعه را نشان میدهد و اول ستون نام نویسنده اول مقاله را بیان میکند.

در قسمتهای بعدی دستورات برنامه **stata** متناسب با این فایل نوشته خواهد شد و شما با فراگیری کار بر روی این فایل میتوانید برای فایلهای دیگر تنها اسم متغیرها را عوض نموده و تحلیلهای مورد نظر را انجام دهید.

### دستور meta

این دستور شاید مهمترین و اولین دستوری باشد که شما برای متا-آنالیز به آن نیاز داشته باشید. این دستور برای شما نتایج مطالعات را ترکیب و مقدار وزن داده شده شاخص مورد نظر را با دامنه اطمینان مربوطه حساب نموده و شاخص مذکور با عدد فرضیه صفر مقایسه میکند. به صورت خودکار برنامه این محاسبات را برای مدل‌های تصادفی و ثابت (**random and fixed models**) تکرار میکند. همچنین آزمون مربوط به عدم تجانس (**heterogeneity test**) نیز به صورت پیش فرض انجام و نتایج آن گزارش خواهد شد. شکل کلی دستور به این صورت است

`meta theta se_theta, gr(f|r) cline xline(value) id(varname)`

در عبارت بالا **theta** یعنی متغیری که مقدار آن برابر شاخص مورد نظر است. متناسب با فرمت فایل بیان شده در ابتدای این فصل نام واقعی متغیر **theta** عبارتند از **logOR**. توجه داشته باشید چون در اینجا از نسبت خطر استفاده میکنید باید مقدار شاخص و انحراف معیار آن

در مقیاس لگاریتمی باشند. ولی برای آنکه جدول و نمودار خروجی در مقیاس لگاریتمی نمایش داده نشوند می‌توان در انتهای دستور عبارت `eform` را اضافه نمود. لذا شکل نهایی دستور به این صورت خواهد بود.

`meta logOR SElogOR, gr(r) cline xline(1) id(trial) eform`

بعد از کاما `gr` نشان می‌دهد که چه نوع نموداری نمایش داده شود. گذاردن `r` و یا `f` در داخل پرانتز به برنامه می‌گوید که نمودار حاصل از تحلیل تصادفی و یا ثابت را نمایش دهد. `cline` به برنامه می‌گوید که خط عمودی بر روی محور عمودی لوزی یعنی مقدار برآورد نقطه‌ای شاخص مورد نظر رسم گردد و `xline(1)` به برنامه می‌گوید که بر روی عدد فرضیه صفر که در اینجا نسبت خطر برابر یک است نیز خط عمودی رسم گردد. عبارت `id(trial)` به برنامه می‌گوید که در نمودار و جدول هر مطالعه را با اسم نویسنده اول مشخص کند و عبارت `eform` نیز همانگونه که در پاراگراف‌های قبل شرح داده شد به برنامه می‌گوید که خروجیها را در مقیاس عددی و نه لگاریتمی به نمایش بگذارد.

خروجی برنامه بعد از اجرای این خط در زیر آورده شده - است:

Method	Pooled Est	95% Lower	CI Upper	Asymptotic z_value	p_value	No. of studies
Fixed	0.672	0.564	0.8	-4.455	0.000	9
Random	0.596	0.4	0.889	-2.537	0.011	

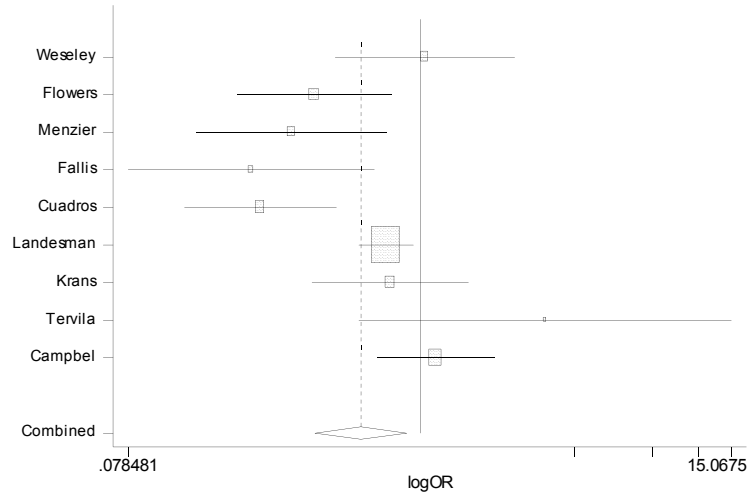
Test for heterogeneity:  $Q=27.265$  on 8 degrees of freedom ( $p=0.001$ )

Moment-based estimate of between studies variance = 0.230

اگر فصل‌های قبل را به دقت مطالعه نموده باشید تفسیر نتایج جدول فوق نباید مشکل باشد. بر اساس نتایج

## جستجوی ساختاریافته و متاآنالیز

آزمون **heterogeneity**، تفاوت بین نتایج مطالعات زیاد می-باشد و بایست نسبت خطر مدل تصادفی را به عنوان جواب آخر دانست که برابر ۰/۵۹۶ بوده و از نظر آماری معنی‌دار است ( $p=0.011$ ). نمودار انباشت این متا-آنالیز نیز به شکل زیر توسط برنامه رسم می‌شود.



البته اگر در دستور فوق بعد از کلمه **print** نیز تایپ شود در خروجی یک جدول دیگر نیز اضافه خواهد شد و اطلاعات تک تک مطالعات را نیز بیان خواهد کرد.

Study	Weights		Study Est	95% CI	
	Fixed	Random		Lower	Upper
Weseley	6.27	2.57	1.04	0.48	2.28
Flowers	8.49	2.88	0.4	0.2	0.78
Menzier	5.62	2.45	0.33	0.14	0.74
Fallis	3.35	1.89	0.23	0.08	0.67
Cuadros	8.75	2.91	0.25	0.13	0.48
Landesman	68.34	4.09	0.74	0.59	0.94
Krans	8.29	2.85	0.77	0.39	1.52
Tervila	1.46	1.09	2.97	0.59	15.07
Campbel	14.73	3.36	1.14	0.69	1.91



در این جدول علاوه بر مقدار نسبت خطر و دامنه اطمینان آن، وزن تعلق گرفته به هر مطالعه در مدل‌های ثابت و تصادفی نمایش داده می‌شود.

بعد از کاما می‌توان عبارت `sort(varname)` را نوشت تا برنامه بر اساس متغیر تعیین شده در داخل پرانتز و به صورت صعودی مطالعات را مرتب و نمودار انباشت مربوطه را رسم کند.

در دستور `meta`، می‌توان بجای مقدار انحراف استاندارد، واریانس و یا دامنه اطمینان را نیز نوشت. در صورتی که بخواهیم واریانس را در دستور بنویسیم، باید به برنامه بگوییم که مقدار گذاشته شده واریانس است نه انحراف استاندارد، برای اینکار باید بعد از کاما عبارت `var` را تایپ نمود. همچنین اگر قبل از کاما و بعد از متغیری که `theta` را نمایش می‌دهد، دامنه اطمینان ۹۵٪ را بنویسیم باید بعد از کاما عبارت `ci` را وارد کنیم که برنامه بداند این اعداد چه می‌باشند.

نکته آخر در مورد این دستور آن است که تنها برای شاخص‌های **Odds Ratio, Risk Ratio, Rate Ratio and Hazard Ratio** باید لگاریتم گرفته شود و مثلاً برای بررسی تفاضل میانگینها نیازی به هیچ اقدامی نیست و به راحتی می‌توان تفاضل میانگینها و انحراف معیار آنها را در دستور وارد نمود.

همچنین اگر بعد از کاما عبارت `ebay` وارد شود در جدول توصیف نتایج `تك تك` مطالعات مقدار شاخص محاسبه شده برای `تك تك` مطالعات به روش بیزین اصلاح خواهد شد (به توضیحات فصل قبل مراجعه شود).

## دستور metan

این دستور نیز همانند دستور قبل عمل می‌کند و خروجیهای آن بسیار شبیه به دستور قبل است. شکل نگارش این دستور برای متغیرهای دو حالت به شکل زیر است

`metan a b c d, options`

که در این دستور `a, b, c, d` به ترتیب خانه‌های جدول دو در دوی زیر است

گروه			
مورد	شاهد		
a	b	+	نتیجه /
c	d	-	پاسخ

شکل دستور `metan` برای متغیرهای عددی (مثلاً تفاضل میانگینها) به شکل زیر است

`metan n1 m1 sd1 n2 m2 sd2, options`

که `n, m and sd` به ترتیب انحراف معیار، میانگین و حجم نمونه در گروههای مورد و شاهد خواهند بود.

بر اساس این دستورات اگر بخواهیم برای فایل مثال دستور `metan` را بنویسیم به این شکل خواهد شد:

`metan p_t p_c np_t np_c, or random`

خروجی این دستور با خروجی دستور `meta` بسیار شباهت دارد و اگرچه ظاهر نمودار انباشت آنها کمی متفاوت است، اولاً اصول کلی آنها یکسان است و ثانیاً با اضافه کردن چند عبارت کوچک می‌توان خروجیهای این دو دستور را کاملاً یکسان نمود.

در این دستور نیز همانند دستور `meta` می‌توان بعد از کاما عبارت `sort(varname)` را نوشت تا برنامه بر اساس متغیر تعیین شده در داخل پرانتز و به صورت صعودی

مطالعات را مرتب و نمودار انباشت مربوطه را رسم کند.

در این دستور بعد از کاما or و random نوشته شده است تا به برنامه بگوید که شاخص مورد نظر Odds Ratio بوده و می‌خواهیم از مدل تصادفی استفاده کنیم. به جای این عبارات می‌توان عبارات زیر را بکار برد

برای متغیرهای دو حالت	
Fixed model using Mantel and Haenszel method	fixed
Fixed model using inverse variance method	fixedi
Random model using Mantel and Haenszel method	random
Random model using inverse variance method	randomi
Peto method in estimation of OR	peto
Odds Ratio	OR
Risk Ratio	RR
Risk difference	RD
برای متغیرهای عددی	
specifies a fixed effect model using the inverse variance method	fixed
specifies a random effects model using the DerSimonian & Laird method	random
computes confidence intervals for odds ratios by Cornfield method	cornfield
pools standardized mean differences by the method of Cohen	Cohen
pools standardized mean differences by the method of Hedges	hedges
pools standardized mean differences by the method of Glass	glass
pools unstandardized mean differences	nostandard

### دستور metacum

این دستور پله به پله اطلاعات مطالعات جدیدتر را با اطلاعات مطالعات قبلی ترکیب می‌نماید و نتیجه را بیان می‌کند. به همین دلیل و همانگونه که در فصل قبل نیز ذکر شده، با اضافه شدن مطالعات و افزایش حجم نمونه اولاً کم کم دامنه اطمینان برآورد پارامتر مورد نظر باریکتر می‌شود و همچنین پایداری آن نیز بالا می‌رود و به عبارتی کمتر از نتایج مطالعات جدیدتر متاثر می‌شود.

برای انجام این تحلیل باید ابتدا با استفاده از دستور `sort` ترتیب مطالعات را بر اساس فاکتور مورد نظر که در اکثر اوقات سال انتشار است مرتب نمود. شکل دستورات به این صورت است:

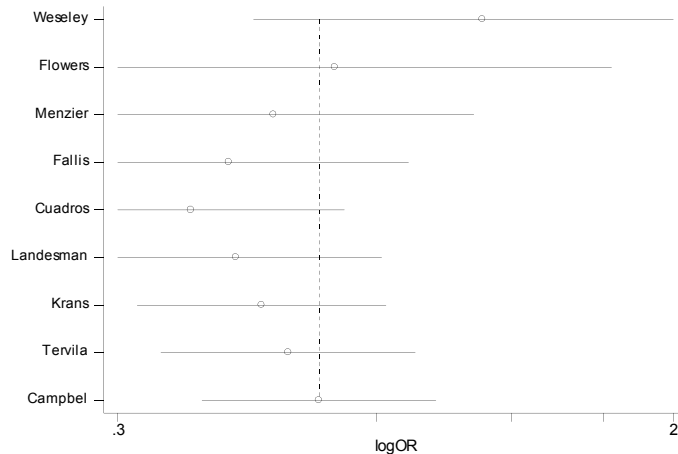
`sort year`

`metacum logOR SElogOR, id(trial) effect(r) eform level(95) graph cline ltrunc(0.3) rtrunc(2)`

در این دستور عبارتهای نا آشنا احتمالاً `ltrunc` and `rtrunc` میباشند. این دو عبارت به برنامه میگویند که محور افقی نمودار انباشت، در سمت چپ و راست به اعداد داخل پرانتز محدود شوند و اگر کمترین و بیشترین اعداد دامنه‌های اطمینان از اعداد فوق فراتر می‌رود، در نمودار ظاهر نشوند. بدین شکل نمودار رسم شده زیباتر جلوه خواهد کرد. این عبارتها را برای نمودارهای دستورات قبلی نیز می‌توان استفاده کرد.

با اجرای این دستور در برنامه `stata` خروجی برنامه به شکل زیر به نمایش در خواهد آمد

Trial	Cumulative estimate	95% CI		z	P-value
		Lower	Upper		
Weseley	1.043	0.477	2.282	0.105	0.917
Flowers	0.63	0.245	1.62	-0.959	0.337
Menzier	0.511	0.258	1.012	-1.926	0.054
Fallis	0.438	0.237	0.809	-2.637	0.008
Cuadros	0.385	0.228	0.651	-3.565	0.0001
Landesman	0.449	0.273	0.739	-3.15	0.002
Krans	0.49	0.321	0.75	-3.289	0.001
Tervila	0.537	0.348	0.829	-2.807	0.005
Campbel	0.596	0.4	0.889	-2.537	0.011



در جدول فوق به ترتیب مطالعات از نظر زمانی مرتب شده‌اند و هر سطر به صورت تجمعی ترکیب اطلاعات هر مطالعه را با مطالعات قبلی به نمایش می‌گذارد. بر اساس این نتایج می‌توان دریافت که با ترکیب نتایج چهار مطالعه اول تقریباً می‌توانستیم به این جمع‌بندی برسیم که مصرف دیورتیکها باعث کاهش خطر بروز پره-اکلامپسی می‌شده است و بقیه مطالعات تاثیر زیادی بر نتیجه‌گیری نهایی نداشته‌اند. همین موضوع در نمودار مربوطه نیز به نمایش گذاشته شده است.

### دستور metabias

این دستور برای بررسی احتمال وجود سوگیر در انتشار نتایج بکار می‌رود و علاوه بر رسم نمودارهای مربوطه، آزمونهای آماری Begg و Egger را انجام می‌دهد.

شکل کلی دستور به صورت زیر است:

**metabias logOR SElogOR, graph(begg|egger) gweight**

عبارت **gweight** به برنامه می‌گوید اندازه نقاط را با توجه به وزن آنها رسم نماید. نتیجه اجرای این دستور به شکل زیر خواهد بود. البته باید در داخل پرانتز یا کلمه **begg** و یا **egger** برای رسم نمودار مورد نظر استفاده نمود.

#### Begg's Test

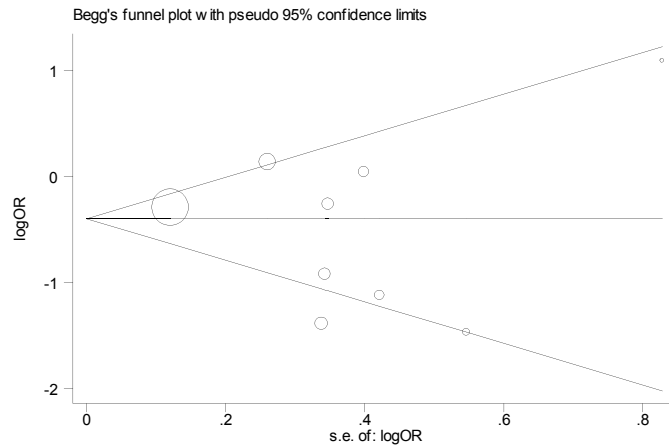
adj. Kendall's Score (P-Q) = -4  
 Std. Dev. of Score = 9.59  
 Number of Studies = 9  
 z = -0.42  
 Pr > |z| = **0.677**  
 z = 0.31 (continuity corrected)  
 Pr > |z| = 0.754 (continuity corrected)

#### Egger's test

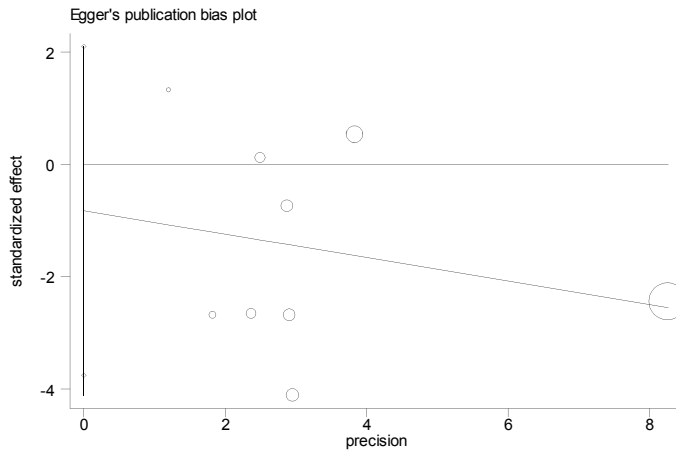
Std_Eff	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
slope	-0.2081	0.331336	-0.63	0.55	-0.99158	0.57539
bias	-0.8271	1.236211	-0.67	<b>0.525</b>	-3.75028	2.096072

بر اساس این نتایج **p-value** وجود سوگیری بر اساس آزمونهای **egger** و **begg** به ترتیب برابر ۰/۶۷۷ و ۰/۵۲۵ می‌باشند. (این اعداد در نتایج فوق مشخص شده‌اند. برای توضیحات بیشتر به فصل هفتم مراجعه نمایید). توجه داشته‌باشید که زمانی لازم است که نتایج **continuity corrected** را گزارش دهیم که در بعضی از مطالعات در یکی از گروههای مورد و یا شاهد تعداد پیامدهای مورد نظر صفر بوده‌باشد. با توجه به اینکه در این متا-آنالیز در تمام گروهها و تمام مطالعات به میزان کافی پره‌اکلامپی دیده‌شده‌است لذا نیازی به این اصلاح وجود ندارد.

در صورتی که در داخل پرانتز عبارت **begg** نوشته شود نمودار کیفی شکل به صورت زیر به نمایش در خواهد آمد که تفسیر کامل آن در فصل هفتم بیان شده است.



و صورتی که در داخل پرانتز عبارت **egger** نوشته شود نمودار مربوطه شکل به صورت زیر به نمایش در خواهد آمد. در این نمودار خط رگرسیون مربوط به تحلیل **Egger** رسم شده و نشان می‌دهد که آیا این خط محور عمودی را در نقطه نزدیک به صفر قطع می‌کند یا نه.



### دستور metatrim

این دستور زمانی کاربرد پیدا می‌کند که آزمونه‌های قبلی نشان داده‌باشند که سوگیری در انتشار نتایج وجود داشته‌است. این دستور با استفاده از روش آماری

Duval and Tweedie nonparametric "trim and fill" method سعی می‌کند تا تخمینی از نتایج مطالعات سانسور شده پیدا نموده و نتیجه متا-آنالیز را تصحیح نماید.

لازم به ذکر است که بسیاری از صاحب‌نظران با انجام چنین تحلیلی موافق نیستند و معتقد هستند که در صورت بروز سوگیری در انتشارات، متدلوژی کار زیر سوال می‌رود و انجام تحلیل و یا تصحیح آن با استفاده از این روش‌های آماری چندان به افزایش دقت کار کمک نمی‌کند. در مقابل بعضی معتقد هستند که وجود سوگیری در انتشارات متاسفانه به صورت بارز و غیر قابل اجتنابی در جستجوی ساختاریافته و متا-آنالیز وجود دارد و رسالت آمار تصحیح این خطاها است و این دستور تا حدود قابل توجه‌ای این رسالت را محقق می‌کند.

البته روش‌های پیچیده‌تری نیز وجود دارد که با استفاده از مدلسازی سعی به رفع خطای سوگیری در انتشارات دارند. تکمیل این این روش‌ها در حال حاضر موضوع روز است و به همین دلیل بحث گسترده در این زمینه در حوصله این کتاب نیست.

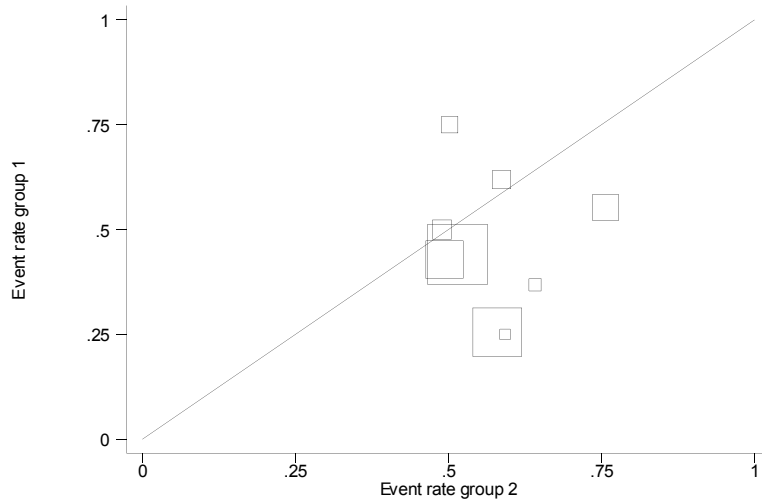
### دستور labbe

این دستور برای رسم نمودار L'Abbe بکار می‌رود و کمک می‌کند تا با استفاده از آن بررسی نماییم که آیا امکان بروز سوگیری در انتشار وجود داشته‌است.



توضیحات کامل این نمودار در فصل هفتم بیان شده است. شکل کلی این دستور به شکل زیر است:

labbe p\_t p\_c np\_t np\_c



### دستور funnel

این دستور نمودار کیفی را رسم می‌کند. از آنجایی که نمودار کیفی به صورت دقیقتر و کاملتر با استفاده از دستور **metabias** رسم می‌شود لذا توصیه می‌شود که از این دستور به صورت مستقیم استفاده نشود. اما به صورت کلی شکل این دستور این گونه نوشته می‌شود.

funnel logOR SElogOR

### دستور metap

این دستور برای انجام متا-آنالیز **p-value** کاربرد دارد و شکل کلی آن به این صورت است

### Metap pvar, e(#)

در این دستور اولین عبارت اسم متغیری را نشان می‌دهد که در بانک اطلاعاتی مقادیر **p-value** را ذخیر نموده است. سپس باید نشان داد که از کدام فرمول برای ترکیب **p-value** ها استفاده شود. به صورت پیش فرض و بدون هر عبارت اضافه‌ای بعد از کاما، فرمول **fisher** استفاده می‌شود. در صورتی که بخواهیم فرمول **Edgington** استفاده شود باید بعد از کاما **e** نوشته‌شود و در داخل پرانتز یا حرف **a** نوشته‌شود تا از متد میانگین استفاده شود (فرمول اول **Edgington**) و یا در داخل پرانتز حرف **n** نوشته‌شود تا از روش توزیع نرمال استفاده شود (فرمول دوم **Edgington**).

(برای توضیحات بیشتر و دانستن کاربرد هر یک از این فرمولها، می‌توانید به فصل قبل مراجعه نمایید)

### دستور metainf

این دستور میزان تاثیر تکتک مطالعات بر روی نتیجه نهایی را ارزیابی نموده و به عبارتی یک تحلیل ساده حساسیت (**sensitivity analysis**) انجام می‌دهد. شکل کلی این دستور عبارت است از:

#### **metainf logOR SElogOR, eform random id(trial) print**

در این دستور از برنامه خواسته می‌شود تا نتایج نهایی از حالت لگاریتم خارج شود (عبارت **eform**)، مدل استفاده شده تصادفی باشد (عبارت **random**)، نتایج هر مطالعه با توجه به اسم نویسنده اول مطالعه اسم گذاری شود (عبارت **id(trial)**) و عبارت **print** به برنامه می‌گوید که نتایج نهایی دستور به صورت جدول ارائه شود.

## جستجوی ساختاریافته و متاآنالیز

با اجرای این دستور جدول زیر توسط برنامه تولید می-شود

<i>Study omitted</i>	$e^{coef.}$	[95% Conf.	
Weseley	0.56	0.36	0.86
Flowers	0.63	0.41	0.97
Menzier	0.64	0.42	0.97
Fallis	0.65	0.43	0.96
Cuadros	0.68	0.47	0.98
Landesman	0.57	0.34	0.97
Krans	0.58	0.37	0.90
Tervila	0.56	0.37	0.82
Campbel	0.54	0.35	0.83
Combined	0.60	0.40	0.89

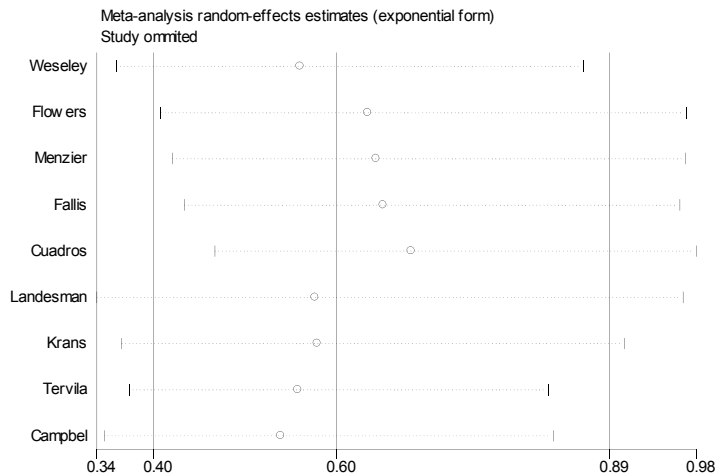
شود

سطر آخر این جدول نشان می‌دهد که در کل نسبت شانس (OR) در کل برابر با  $0/6$  بوده است. در سطرهای بالاتر نشان می‌دهد که با حذف هر مطالعه میزان نسبت شانس بدست آمده چه مقدار خواهد بود. همانگونه که مشاهده می‌شود مثلاً حذف اولین مطالعه (Weseley) باعث خواهد شد که نسبت شانس محاسبه شده به  $0/56$  برسد و عملاً می‌توان گفت که این مطالعه به میزان حدود  $0/04$  این عدد را تغییر داده است. بر اساس نتایج فوق مشخص است که بیشترین تاثیر را مطالعه Cuadros داشته است و باعث شده است که به میزان  $0/08$  نسبت شانس به تغییر کند.

با اجرای دستور مذکور برنامه همین اطلاعات را به صورت یک نمودار نیز نمایش می‌دهد. اگرچه ظاهر این نمودار شبیه نمودار انباشت است ولی مفهوم آن متفاوت است و هر سطر مقدار نسبت شانس جمع‌ی را بدون وارد

## جستجوی ساختاریافته و متاآنالیز

نمودن اطلاعات مطالعه مذکور را نمایش می‌دهد. در این نمودار سه خط عمودی نمایش داده شده دامنه اطمینان و مقدار نسبت شانس کل بر اساس اطلاعات تمامی مطالعات را نشان می‌دهند.



### دستور metatrim

این دستور سعی می‌کند مقدار سوگیری در انتشار را کمی نموده و با استفاده از روش Duval and Tweedie nonparametric "trim and fill" مقادیر متناظری برای اطلاعات مطالعات ایجاد کننده سوگیری پیش‌بینی نموده و نتایج نهایی را تعدیل کند. شکل کلی دستور به این شکل است:

**metatrim logOR SElogOR, reffect eform print funnel**

عبارت **reffect** به برنامه می‌گوید که مدل مورد استفاده تصادفی باشد. عبارت **eform** باعث می‌شود که نتیجه نهایی از حالت لگاریتمی خارج شود و با آوردن عبارت **print** از برنامه خواسته می‌شود که نتایج تحلیل در خروجی **stata** نمایش داده شود. در انتها دستور نیز عبارت **funnel** به

برنامه می‌گوید که یک نمودار کیفی ترسیم نماید که در آن علاوه بر اطلاعات مربوط به تک تک مطالعات وارد شده در تحقیق، اطلاعات نقاطی که به نظر در اثر سوگیری یافت نشده اند را نیز به نمایش بگذارد.

از آنجایی که در این فایل، سوگیری در انتشار نتایج با توجه به نتایج آزمونهای آماری Egger و Begg وجود نداشته است لذا این دستور نیز تغییر قابل ملاحظه‌ای در نتیجه‌گیری ایجاد نخواهد کرد. ولی در صورتی که در اطلاعات سوگیری وجود داشته‌باشد، نتیجه بدست آمده در این دستور با نتایج بدست آمده در اثر اجرای دستورات ساده مانند meta تفاوت قابل ملاحظه‌ای ایجاد خواهد شد.

---

۱۰. منابع

---

1. **Stata User's Guide.** College Station Texas: Stata Press.
2. **Stata Reference Manual.** In. Texas: College Station, Stata Press; 2001.
3. **Intercooled STATA 8.0 for Windows.** Texas: College Station, Stata Corporation; 2003.
4. Bailar JC: **The Promise and Problems of Meta-Analysis.** *N Engl J Med* 1997, **337**:559-561.
5. Cleves M, Gould WW, Gutierrez RG: **An Introduction to Survival Analysis Using Stata, Revised Edition.** College Station Texas: Stata Corp; 2004.
6. DerSimonian R, Laird N: **Meta-analysis in clinical trials.** *Control Clin Trials* 1986, **7**(3):177-188.
7. Duval S, Tweedie R: **Trim and Fill: A Simple Funnel-Plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis.** *Biometrics* 2000, **56**(2):455-463.
8. Duval S, Tweedie R: **A Nonparametric" Trim and Fill" Method of Accounting for Publication Bias in Meta-Analysis.** *Journal of the American Statistical Association* 2000, **95**(449).
9. Eddy DM, Hasselblad V, Shachter R: **An introduction to a Bayesian method for meta-analysis: The confidence profile method.** *Med Decis Making* 1990, **10**(1):15-23.
10. Egger M, Smith GD: **Meta-analysis: Potentials and promise.** *British Medical Journal* 1997, **315**(7119):1371.

11. Egger M, Smith GD: **meta-analysis bias in location and selection of studies**. *BMJ* 1998, **316**:61-66.
12. Egger M, Smith GD, Phillips AN: **Meta-analysis: Principles and procedures**. *British Medical Journal* 1997, **315**(7121):1533.
13. Fleiss JL: **The statistical basis of meta-analysis**. *Stat Methods Med Res* 1993, **2**(2):121-145.
14. Galbraith RF: **A note on graphical presentation of estimated odds ratios from several clinical trials**. *Stat Med* 1988, **7**:889-894.
15. Gould W, Pitblado J, Sribney W: **Maximum Likelihood Estimation with Stata**. Texas: Stata Press; 2006.
16. Hardy RJ, Thompson SG: **Detecting and describing heterogeneity in meta-analysis**. *Statistics in Medicine* 1998, **17**(8):841-856.
17. Higgins JPT, Thompson SG: **Quantifying heterogeneity in a meta-analysis**. *Statistics in Medicine* 2002, **21**(11):1539-1558.
18. Hunter JE, Schmidt FL: **Methods of Meta-Analysis: Correcting Error and Bias in Research Findings**. New York: Sage Publications; 2004.
19. In RA: **The Hazards of Scoring the Quality of Clinical Trials for Meta-analysis**. *JAMA* 1999, **282**:1054-1060.
20. J Sweeting M, J Sutton A, C Lambert P: **What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data**. *Statistics in Medicine* 2004, **23**(9):1351-1375.
21. Lindbaek M, Hjortdahl P: **How do two meta-analyses of similar data reach opposite conclusions**. *BMJ* 1999,

318(7187):873-874.

22. Lipsey MW, Lipsey M, Wilson DB: **Practical Meta-Analysis**: Sage Publications; 2001.
23. Macaskill P, Walter SD, Irwig L: **A comparison of methods to detect publication bias in meta-analysis**. *Statistics in Medicine* 2001, **20**(4):641-654.
24. Normand ST: **Tutorial in biostatistics meta-analysis: formulating, evaluating, combining, and reporting**. *Statistics in Medicine* 1999, **18**:321–359.
25. Petitti DB: **Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis**. New York: Oxford University Press; 1994.
26. Pogue J, Yusuf S: **Overcoming the limitations of current meta-analysis of randomised controlled trials**. *Lancet* 1998, **351**(9095):47-52.
27. Raudenbush SW, Bryk AS: **Hierarchical Linear Models: Applications and Data Analysis Methods**: Sage Publications; 2002.
28. Rosenthal R, DiMatteo MR: **META-ANALYSIS: Recent Developments in Quantitative Methods for Literature Reviews**. *Annual Review of Psychology* 2001, **52**(1):59-82.
29. Smith TC, Spiegelhalter DJ, Thomas A: **Bayesian approaches to random-effects meta-analysis: a comparative study**. *Stat Med* 1995, **14**(24):2685-2699.
30. Sterne JA, Egger M: **Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis**. *J Clin Epidemiol* 2001, **54**(10):1046-1055.
31. Sterne JAC, Gavaghan D, Egger M: **Publication and related bias in meta-analysis: power of statistical tests**



- and prevalence in the literature.** *J Clin Epidemiol* 2000, **53**(1119):29.
32. Stewart LA, Clarke MJ: **Practical methodology of meta-analyses (overviews) using updated individual patient data.** *Cochrane Working Group. Stat Med* 1995, **14**(19):2057-2079.
33. Stewart LA, Parmar MK: **Meta-analysis of the literature or of individual patient data: is there a difference?** *Lancet* 1993, **341**(8842):418-422.
34. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB: **Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group.** *JAMA* 2000, **283**(15):2008-2012.
35. Sutton AJ: **Methods for meta-analysis in medical research.** London: Wiley Chichester; 2000.
36. Tang JL, Liu JL: **Misleading funnel plot for detection of bias in meta-analysis.** *J Clin Epidemiol* 2000, **53**(5):477-484.
37. Thompson SG: **Why source of heterogeneity in meta-analysis should be investigated.** *BMJ* 1994, **309**:1351-1355.
38. Thompson SG, Pocock SJ: **Can meta-analysis be trusted.** *BMJ* 1991, **339**:1127-1130.
39. Thompson SG, Sharp SJ: **Explaining heterogeneity in meta-analysis: a comparison of methods.** *Statistics in Medicine* 1999, **18**(20):2693-2708.
40. Tobias A: **meta-analysis of p-values.** *Stata Technical Bulletin* 1998, **49**:15-17.

41. Wolf FM: **Meta-Analysis: Quantitative Methods for Research Synthesis**. New York: Sage Publications; 1986.